

What to do about European union

That Denmark should have thrown dust in the eyes of those working for a politically united Europe only illustrates the force of last week's argument that Europe needs ways of making decisions that command general respect.

Who would have thought that 50,000 voters in last week's Danish referendum on the Maastricht Treaty would so easily have confused the European enterprise? Danish commentators had been saying in advance that the result would be close, but had assumed that calculations of the economic benefits of continued membership of the European Communities (EC) would, in the end, prevail. In the event, they did not, which is no great surprise. Uncompliant Danes could well have concluded that to vote against the treaty would leave Europe exactly as it is, without further supranational integration and without changing Denmark's status. The other 11 member governments have responded by putting their heads firmly in the sand, saying that they will press ahead with integration, hoping that Denmark will change its mind before the year is out. Meanwhile, lawyers in Brussels are touting a scheme whereby the surviving members would renounce the Treaty of Rome, and adopt Maastricht instead. The snag there is that there would have to be a rerun of the whole ratification process with a different treaty.

The moral in this muddled tale is simply that it is yet another, if almost melodramatic, proof that Europe's way of reaching decisions is cack-handed (see *Nature* 357, 347; 4 June 1992). So much should be clear from the way in which the final version of the treaty was put together last December, by horse-trading between heads of government. Mr John Major's supposed victories on behalf of Britain (that the planned move to a single currency in four years would not apply if Britain then chose otherwise, and that a raft of social legislation would be optional) now seem hollow, national derogations from a treaty that may no longer have force. What now emerges is that horse-trading may satisfy those who participate, but is not more widely persuasive.

What, in this crisis for the EC, should be done? The plain truth is that most member governments are lumbered with a substantial (if minority) army of domestic sceptics. Many of them, including both the British and the German governments, would not wish to follow Denmark in a referendum, but would rather hurry on with ratification. The President of France, on the other hand, has been forced to promise to take the risk this autumn, mostly out of fear that next year's elections to the national assembly would otherwise be compromised. The most obvious difficulty, then and now, is that the people of Europe know too little about the treaty and its implications, but too much for their own enlightenment of the opinions of those who are against it.

That is why the most urgent need is for a programme of

public education. Governments are reluctant to take on this task for fear of exciting domestic opposition to further European integration. But is it not silly of them to hope that their own electors will joyfully sign on for the replacement of national currencies by a single European currency, for example, when the advantages of the change are so obscure, and the disadvantages (and inconveniences) so easily magnified and trumpeted? The difficulty is that the governments are unlikely to change. That is why the best course would be that the European Commission should take on the job of making sure that Europe as a whole understands what the European Union (for which Maastricht would legislate) is about. It would be a gigantic task, requiring that the Maastricht Treaty should be unpackaged into its components, each of which would have to be explained with care. But it is a job worth doing for its own sake. And quickly.

But that can be only a stop-gap. For the longer haul, there needs to be a way of making the proposed European Union more democratic than is now intended. It will not have escaped the attention of the officials in Brussels and the member governments that they are in danger of being overwhelmed by populism (represented by people as different as Mr Ross Perot from Texas, who would become the president of the United States, and Mr Vladimir Meciar, who would make Slovakia separate from Czechoslovakia). The convention that national governments, for the duration of their legal terms, embody the sovereign will of their electors is no longer widely acceptable. In any case, governments need not merely the dull acquiescence of their electors if they are to stay in office, but their active consent as well. The case for a more democratic European Union than that now foreseen is too urgent to be left for the review of Maastricht now planned for 1996. □

Sound embryo ruling

A Tennessee court accepts that an eight-cell blastomere is not advanced enough to be considered a person.

How are the thousands of human embryos at the four-to-eight cell stage of division, created *in vitro* from the eggs and sperm of men and women unable to conceive naturally and now frozen in liquid nitrogen in laboratories throughout the United States, to be disposed of? That question and others



were tackled intelligently two weeks ago by the Supreme Court of Tennessee. This is how the argument goes. If the genetic parents of frozen embryos wish that one or more should be implanted in the uterus of the woman, physicians and medical ethicists see no problem. After all, the *in vitro* fertilization was presumably undertaken to help the parents give birth to a child. But what if one of the genetic parents dies, leaving fertilized gametes unused? Or what if a couple divorce and cannot agree on the disposition of their potential progeny? That is what happened in the case of Junior Lewis Davis and Mary Sue Davis (now Mary Sue Stowe), whose contentious divorce included a dispute about the disposal of seven frozen embryos the couple had created in happier times.

Mary Sue Stowe, no longer wishing to have Junior Davis's child, wants to donate the embryos anonymously to an infertile couple. Junior Davis, arguing that he does not wish to be the anonymous father of someone else's child, would like to see the embryos discarded. On 1 June, Justice Marth Craig Daughtrey of the Supreme Court of Tennessee issued a Solomon-like decision in the case, which had been heard previously in two lower state courts. Her opinion, which is an extraordinarily lucid exposition of the issues, hangs on a person's right to privacy as implied, in her reading, by both the US and state constitutions. The court ruled that Junior Davis has a right not to become a parent against his will.

The most interesting feature of the decision is the reasoning behind the court's conclusion that a frozen embryo at the four-to-eight cell stage is legally a "preembryo", not an embryo at all. "[S]emantical distinctions are significant in this context", Justice Daughtrey wrote, "because language defines legal status and can limit legal rights." In defining a preembryo, the court relied on scientific data regarding fertilization and development in a 1990 report of the American Fertility Society, *Ethical Considerations of the New Reproductive Technologies*.

The society, represented in the proceedings *amicus curiae*, said of the embryo after three divisions that: "Each blastomere, if separated from the others, has the potential to develop into a complete adult. Stated another way, at the 8-cell stage, the developmental singleness of one person has not been established." As cell division continues, the outer and inner cells of the developing organism become increasingly different, particularly as the outer cells of the blastocyst implant into the uterine wall. "Thus, the first cellular differentiation of the new generation relates to physiological interaction with the mother, rather than the establishment of the embryo itself."

In accepting this description of the beginning of life, the Tennessee court explicitly rejected a lower court ruling, based on the testimony of French geneticist Jerome Lejeune that "life begins at the moment of conception" and that four-to-eight cell entities are "tiny persons" and thus his "kin". By defining a preembryo as it did, the Tennessee Supreme Court rejected both the proposition that a frozen embryo is a person, with full legal rights, and that an embryo is simply property that can be divided like pieces of silver, but said,

instead, that preembryos "occupy an interim category that entitles them to special respect because of their potential for human life."

Recognizing both that the Davis case is probably without precedent and that *in vitro* fertilization is a growing enterprise, the court also went beyond the strict legal limits of the case to offer advice in other ethically difficult areas. Thus it extended the definition of parenthood to include genetic parenthood (thereby honouring Junior Davis's privacy right not to become a father) and also stated that a preembryo is not entitled to protection under laws embodying a state's legitimate interest in protecting life. There are also valuable practical suggestions; men and women undertaking *in vitro* fertilization should carefully consider the disposal of unwanted preembryos before the procedure, and then bind themselves to donate or discard the embryos or contribute them to research. The genetic parents, who enjoy the rights, must accept the responsibility.

By weighing the issues as thoroughly as it has done, the Tennessee Supreme Court has done a service to physicians, researchers and potential parents. It is to be hoped that its arguments will be widely accepted. Perhaps the most valuable outcome of the decision would be a general understanding, among those who oppose contraception in particular, that the definition of life is not as simple as it is sometimes made to seem. □

Mrs Brundtland's wish

The prime minister of Norway, having popularized sustainable development, now wants world government.

Mrs Gro Harlem Brundtland, who appeared at Harvard University's commencement ceremonies on 4 June straight from a red-eye flight from Rio de Janeiro, is an admirable and admirably forceful woman. Like many others, she wants to make the world a better, even a much better, place. She has won a great deal of what she wants already: the current conference at Rio is happening partly because of the urging of the international commission of which she is a member, for example. She explained last week her disappointment that the conference will say nothing about population growth, which will be widely shared.

But there is also a sense in which Brundtland is asking for the Moon. Last week she argued cogently that the Gulf War was an example of supranational action against an aggressor. She went on, correctly, to argue that many problems in development and environment are supranational, but then went on to conclude that their solution requires a measure of world government. That, too, is a widely shared ambition, but is it attainable in the short term? The difficulties of creating a European government (see above) have given pause to the European enterprise: can it be in the best interests of development and the environment that their amelioration should wait on world government? Do not the recent successes in arms control suggest that urgent hatchets can be buried without waiting for the abolition of sovereignty? □

NIH laboratory admits to fabricated embryo research, retracts paper

Washington. As retractions go, the one in the 29 May issue of *Cell* was short and unambiguous: data supporting a 1991 paper on mouse embryo development written by four researchers from the US National Institutes of Health (NIH) had "been fabricated by one of the authors ... without any knowledge by the others".

But cases of scientific misconduct these days are rarely that simple. Although there is no argument about what was done and who did it, the NIH case raises troubling questions about the ability of researchers in a laboratory to block a determined effort by a colleague to fabricate data.

The paper, entitled "Oct-3 is a maternal factor required for the first mouse embryonic division" (*Cell* 64, 1103; 1991), came to the surprising and unexpected conclusion that embryonic cell division requires Oct-3, a protein that switches on genes, and that the protein may even regulate DNA replication in early development. This was more than anyone had expected of a humble transcription factor.

Oct-3, as it turns out, does nothing of the sort. When Louis Staudt and other researchers in his laboratory at NIH's National Cancer Institute (NCI) tried to replicate the results, they found that the key experiments no longer worked. Yet the only significant difference between their work and what had appeared in *Cell* was the people doing it. The first author was Mitchell Rosner, a graduate student who spent last year at NIH but who is now at the Harvard University Medical School.

When Rosner learned that Staudt could not replicate his findings, he volunteered to repeat the experiments. Staudt and other researchers in the laboratory set up a double blind test. Rosner prepared his reagents as usual, and when the codes were revealed the experiment again worked. But this time, when his colleagues examined the reagents, they discovered why.

Initially, contamination in the injected samples interfered with the embryonic development, a problem that had apparently disappeared after Rosner refined the reagents. In the subsequent investigation, the principal samples were shown to contain just what they should — antisense Oct-3 oligonucleotide to block Oct-3 in the developing embryos.

But the controls, rather than containing DNA without antisense Oct-3, turned out to be simply a buffer solution — essentially water. Rosner had 'solved' the contamination problem by removing the DNA entirely. In the replication trial he had apparently broken the code and, as he had in the

previous experiments, surreptitiously switched test tubes for the control groups. No wonder, then, that the control embryos performed nearly as well as those that were not injected at all.

Confronted with his colleagues' suspicions, Rosner confessed. Staudt says that Rosner blamed the fabrication on "self-imposed pressure". Rosner did not return telephone calls last week.

Staudt reported the misconduct to officials at NCI and NIH. Philip Chen, the acting director of the NIH intramural programme, confirms that the case has been given to the NIH Office of Scientific

***"Recent investigations have revealed that the experimental evidence supporting the conclusions of the paper by Rosner et al. has been fabricated by one of the authors (M.R.) without any knowledge by the others. We therefore retract this paper in its entirety. We sincerely apologize to anybody, within or outside the research community, who has been misled by this publication."* — Cell.**

Integrity, which is in the midst of an extensive reorganization (see page 431).

Suzanne Rouffembart, a spokeswoman for Harvard Medical School, says that Rosner told the school about the case and that the dean has assembled an advisory committee of faculty members. The committee is expected to submit a report later this month. Rosner, who was a candidate for a doctoral degree at Georgetown University Medical School when the paper appeared, has also withdrawn his candidacy for that degree.

Researchers familiar with the case say that it illustrates both the length to which a researcher may go to hide data fabrication and the difficulty of preventing such behaviour.

"If you didn't know it was a hoax, you'd say it was a fine piece of work, a well-controlled study," says Michael Green of the University of Massachusetts Medical Center. Confronting such deliberate fabrication, he says, "is frightening, but when the fright passes, you realize that it's virtually impossible to stop a person who is willing to act in a determined and malicious fashion."

By all accounts, the Rosner case did not fit either of two stereotypes of scientific misconduct: he was neither a junior researcher working alone and submitting data to a distracted laboratory chief nor an un-

known student producing amazing results overnight for supervisors too pleased to examine the data closely.

Staudt says he worked closely with Rosner on the experiment and reviewed his progress daily. "I saw him put in the DNA tablets," he says. (Rosner apparently switched test tubes when he took the controls to part of the experiment in another room.)

In addition, the results were the product of long, hard work, and were built upon earlier efforts. Rosner was the lead author of a paper in *Nature* (354, 686; 1990) that Staudt says "took us a year to do and has been proven entirely correct and confirmed by other labs." Furthermore, Staudt says, Rosner was "working day and night".

To Alan Rabson, director of the Division of Cancer Biology, Diagnosis and Centers at NCI, "this case points out that a truly determined person can deceive even a full-time, conscientious scientist". To others, the episode proves that scientists can indeed police themselves. "This is a case of science working as it should," says Benjamin Lewin, the editor of *Cell*.

Staudt says that the case has taught him that "fraud is present and needs to be watched out for. I wish I had done the experiment again, with other people [preparing the reagents]. In the future, I'll be more ready to do that."

Others, however, doubt that anyone can deter an inventive researcher determined to fabricate data. They point out that replicating an experiment goes beyond repeating a single procedure and could require asking an outside researcher to recreate reagents that have taken months to produce. And they say that the extra effort is not justified.

"Even if you could find an independent person in your lab who would be willing to do the whole experiment again", says Green, "that would reduce your scientific productivity and increase the cost of the science by as much as a factor of two. The price you'd pay for that is not worth stopping the .001 per cent of fraudulent manuscripts."

The 15-month delay between publication and retraction was time enough for dozens of unsuspecting scientists to waste their time chasing the erroneous implications of the paper. Although that does not appear to have happened in this case, the potential damage from fraudulent science entering the literature has convinced Staudt that extra precautions, no matter how inconvenient, are needed to avoid another case of misconduct. That lesson, however, appears to be something that can be learned only at first hand.

Christopher Anderson

Canada commits money for human genome research

Quebec. The Canadian government has decided to spend C\$22 million (US\$18 million) over the next five years on research to map and sequence the human genome.

Speaking at an international industrial biotechnology conference in Montreal on 2 June, Science Minister William Winegard and Health Minister Benoit Bouchard said C\$17 million will come from the federal government — C\$12 million of it from the Industry, Science and Technology Canada department (ISTC) and C\$5 million from the Medical Research Council of Canada (MRC). The remainder will come from the private National Cancer Institute of Canada (NCIC). The ISTC contribution is new money, while MRC's and NCIC's will come from their current budgets. NCIC's budget is derived from public donations to the Canadian Cancer Society.

The announcement is a victory for a group of prominent biomedical scientists who have tried to persuade the Canadian government to invest in such a project. "Four years ago we got a nice reception from the minister [Winegard], who was very sympathetic, but we were told there wasn't any money", said Louis Siminovitch, a medical geneticist who is director of the Samuel Lunenfeld Research Institute of Toronto's Mount Sinai Hospital.

Asked what might have changed the government's mind, Henry Friesen, MRC's new president, said he thinks that officials decided that it would be a mistake not to take part in the programme. "Maybe the case wasn't made as strongly in the past", he said. "There still is in some quarters a scepticism whether spending on this scale is justified. The resources here are relatively modest. But I think there's also growing recognition of what we're going to learn, and that it may cost us a lot more if we're not part of the club."

US genome officials are pleased that their northern neighbour is investing in research on the human genome. "We're very much looking forward to collaborating with Canada. It sounds very encouraging", says Elke Jordon, deputy director of the Office of Human Genome Research at the US National Institutes of Health (NIH). "We have been expecting this for some time, but I know there were delays."

The Canadian prime minister, Brian Mulroney, has been willing to provide money for projects in which the private sector shares the cost. Friesen's new long-term strategic plan for MRC emphasizes partnerships (see *Nature* 357, 271; 1992), and in talks with government officials before his appointment Friesen used the genome project as a case in point. The Pharmaceutical Manufac-

turers' Association of Canada has also expressed an interest in the genome effort.

ISTC's contribution to the project, and the fact that it was announced at an industrial conference, suggests that the government sees the genome project as economically profitable. "This is the kind of investment that governments must undertake in partnership with the stakeholder [and that] contributes to ... competitiveness and to our nation's long-term prosperity", says Winegard, who referred to applications "ranging from health care to agriculture, industrial biotechnology, environmental remediation, forestry and informatics".

NCIC's unsolicited contribution of C\$5 million to the genome project is a reflection of the important role that genes play in cancer research. "This was the hot new topic that fired everyone's imagination", says Friesen, a former NCIC president.

The kinds of research to be carried out and the balance between biology and technology will be decided by a management team headed by Ronald Worton, professor of medical genetics at the University of Toronto and geneticist-in-chief at the Hospital for Sick Children. Worton directed research that led to the identification and characterization of the gene responsible for Duchenne muscular dystrophy.

He is also one of 11 Canadian members of the international Human Genome Organization (HUGO) and a member of the advisory committee that made recommendations to Canada's three research funding councils on the formation of the Canadian programme. These are the MRC, the Natural Sciences and Engineering Research Council (NSERC) and the Social Sciences and Humanities Research Council (SSHRC). Together with ISTC and NCIC, they make up the genome programme management committee.

Friesen thinks that the project will fund both research proposed by the advisory committee and investigator-initiated proposals. He guesses that the latter "would be the dominant activity". He expects that both the human genome and model species will be explored. "There's considerable strength in mouse genetics and yeast genetics in Canada", he says.

The genome project should also stimulate the development of new technology, possibly in collaboration with industry. The project will provide money for training programmes, in particular for postdoctoral fellows, and 7.5 per cent of its funds will be devoted to the social, ethical and legal issues of genetic research. The first round of grants will be awarded in late autumn after a review in September.

David Spurgeon

Industry surprised by firm US stance on biodiversity treaty

Washington. When US President George Bush flatly refused last week to negotiate a biodiversity treaty at the Earth Summit because of what he saw as an unacceptable tradeoff between biological diversity and US jobs, he disappointed environmentalists and made headlines. He also surprised the very industry he was claiming to defend.

Officials at two major biotechnology associations say they do not oppose the gist of the treaty. Their concerns, expressed in letters to the White House, focused on draft language that could set a precedent for government control over industry licensing and technology transfer practices. But they had hoped that the offending language could be modified rather than made moot by the president's unyielding stance.

Bush refused to negotiate, however, and even rejected an appeal by William Reilly, head of the US Environmental Protection Agency, to consider compromise language suggested by the Brazilians. "I have to be the one ... responsible for jobs," Bush told reporters last week. "I have nothing to be apologetic for."

Contrary to the impression left by Bush, the biotechnology industry was not wringing its hands over the proposed treaty. On 18 May, the Association of Biotechnology Companies (ABC) wrote a letter to Bush devoted largely to its support of a controversial gene patent application by the National Institutes of Health. It attached an eight-page position paper on the subject. ABC also said in the letter that it would oppose "any provision that would eliminate or dilute the protection of intellectual property rights for biotechnology ... or would require or permit national legislation compelling the transfer of technology". The accompanying position paper was two paragraphs.

William Sweet, president of the ABC, says he was "kind of surprised" by the strong US response. The Industrial Biotechnology Association (IBA) held a meeting with White House officials in what Lisa Rains, IBA's vice president for government relations, calls a "low-keyed lobbying effort" to modify the proposed treaty.

Pharmaceutical companies, another group likely to be affected by the intellectual property provisions of the treaty, were caught off-guard by the US position. The Pharmaceutical Manufacturers Association did not even discuss the treaty until this week, according to a spokesman. "It's puzzling that a number of other nations with a substantial interest in the technology have somehow made themselves comfortable enough to sign the treaty", says James Russo, director of government and public policy for Smith-Kline Beecham.

Christopher Anderson

CERN officials visit Tokyo, hat in hand, in search of funds for their next big collider

Tokyo. These days, all roads to expensive high-energy physics projects seem to go through Japan.

Officials from the European Laboratory for Particle Physics (CERN) turned up last week in their first-ever attempt to woo the Japanese. What they were selling is a piece of the laboratory's next big collider project, the Large Hadron Collider (LHC), which they hope to turn on before the end of the decade. Walter Hoogland, director of research at CERN, and James Allaby, leader of CERN's experimental physics division, toured all of Japan's science-related ministries and agencies as part of a campaign to find outside support for the \$1,500-million machine.

The visit comes in the midst of a determined effort by US officials to persuade the Japanese to help pay for their next big machine, the \$8,500-million Superconducting Super Collider (SSC), also scheduled to begin work in 1999. Earlier this year, the United States and Japan formed a Joint Working Group to discuss Japan's contribution to the proton-antiproton accelerator, and SSC officials are confident (see story below) that they have the inside track on whatever money the Japanese decide to spend

on high-energy physics projects beyond their shores.

Hoogland and Allaby did not specifically ask the Japanese for financial support to build the LHC. That is most likely to happen in August, when Carlo Rubbia, CERN's director general, visits Tokyo. But Japan has a limited amount of money for such international collaborations, and it probably cannot afford to make a significant contribution to both machines.

The CERN officials and European embassy officials in Tokyo deny that they are competing with the SSC laboratory for Japanese funds. The two projects, they say, are "complementary". But in their discussions with the Japanese they made no secret of the advantages they think the LHC and CERN hold over the SSC.

Speaking at a reception at the British Embassy, Hoogland pointed out that CERN projects are less expensive because they use existing accelerators as a "stepping stone" to the next machine. In the case of the LHC, that means using the tunnel built for the Large Electron-Positron (LEP) accelerator. The European collider, he points out, can draw upon a well-tested infrastructure and an experienced technical staff, whereas the

SSC has to be built and tested "from scratch". CERN, with 17 members, also has much more experience with large international collaborations, the Europeans add.

Those officials say that it is because of the international character of CERN that they are now knocking on Japan's door. In the past few years, the number of Japanese scientists using CERN's facilities has grown to 70, compared with only a dozen in 1982. This presence is comparable to the number of users from the Netherlands, which is a member of CERN. But, unlike the Netherlands, Japan pays nothing for this access, and CERN officials feel the time has come for some financial contribution.

CERN hopes that non-member nations that use CERN's facilities will contribute roughly 20 per cent, or about \$300 million, of the cost of building the LHC. The US Department of Energy has promised Congress that one-third of the initial cost of the SSC, or about \$2,800 million, will come from non-federal sources; Texas has already pledged \$1,000 million, leaving some \$1,800 million to come from the rest of the world.

It is too early to predict how the Japanese will respond to CERN's entreaties. An official at the Ministry of International Trade and Industry (MITI) who met Hoogland and Allaby says that "we don't think about the LHC, only the SSC". However, a foreign ministry official says that "both projects have merit" and that the Japanese government will consult Japanese scientists for advice.

Many Japanese high-energy physicists say that the decision on whether to contribute to the SSC has become a matter of politics, not science. The decision will be up to Kiichi Miyazawa, the prime minister, who has been under considerable pressure from the US president, George Bush.

Although no one pretends to know what Miyazawa will do, one official from Japan's Science and Technology Agency says that the timing of the European approach is "very good". The Japanese cabinet, in line with numerous recommendations from government committees, recently called for a doubling of the government budget for science. But it did not specify a timeframe. A vigorous debate is now going on between the Ministry of Finance and a science pressure group within the ruling Liberal Democratic Party on when and how to implement the increases. Most of the extra money is expected to go to domestic research, but some may be available for international programmes.

David Swinbanks

SSC says it holds the edge

Washington. Roy Schwitters, director of the Superconducting Super Collider (SSC) laboratory, believes that the United States has built up the sort of diplomatic ties with Japan that are needed to win any financial commitment from the Japanese for the proton-antiproton collider. That is not true for his European colleagues from CERN, he says, who are hoping that the Japanese will also help pay for the laboratory's proposed Large Hadron Collider.

"You shouldn't underestimate the value of the Joint Working Group that we have established", says Schwitters, referring to the administrative arrangement created in January when the US president, George Bush, met the Japanese prime minister, Kiichi Miyazawa, in Tokyo. "The Japanese put great weight behind the existence of that group. To them, it signifies how committed they already are to the project." The group met in Tokyo for the first time in April, and has formed panels that will study specific issues and report back to the working group in the autumn.

The delegation from CERN, on the other hand, has no official status in Japan, Schwitters points out. And without it, he says, it is all but impossible to turn even the best intentions of scientists and bureaucrats into hard cash.

"Carlo Rubbia [CERN's director general] is a very respected senior scientist, but he doesn't have the benefit of our formal relationship", says Schwitters. "And believe me, that makes a big difference."

Meanwhile, Schwitters has problems of his own in shepherding the SSC through the US political process. Last week, for example, the energy and water appropriations subcommittee in the US House of Representatives slashed the president's 1993 request for \$650 million for the SSC, recommending instead that the laboratory receive only its current budget of \$484 million.

"That will mean delays", says Schwitters about the first of several congressional votes this summer on the project. "It's very hard to ramp up a project when you don't get what you need. In effect, level funding would force us to start cutting back just when we should be moving forward."

Jeffrey Mervis

Gemini telescope project shifts into high gear

Washington. With a lot to do and not much time, an international collaboration to build two 8-metre telescopes by the end of the decade has put Sidney Wolff temporarily in charge of the \$176-million project.

Wolff has taken a six months' leave from her position as director of the National Optical Astronomy Observatories (NOAO), which operates three major facilities with funding from the US National Science Foundation. The telescopes, known as the Gemini Project because of twin sites in Hawaii and Chile, were originally intended to be funded entirely by the United States. But in the past three years the project has become international, with the United States putting up half the total, Britain paying 25 per cent and Canada 15 per cent. The remaining 10 per cent is being sought from one or more countries in Latin America, including the host country of Chile.

Wolff was enlisted as a troubleshooter precisely because of the additional work required to pull off such an arrangement. "Being an international collaboration makes it much more complicated", says Bob Bless, an astronomer at the University of Wisconsin and a member of the project's eight-member international advisory board. "A lot of things have to happen in the next few months, and it seemed to be more efficient to let the technical people do their thing and get somebody else to handle the administrative aspects. Sidney is already

familiar with the project, and she's also a terrific administrator."

Wolff says that she must work out the financial arrangements as soon as possible so that scientists can know how much money is available for the two telescopes. The nature of the collaboration will be spelled out in a memorandum of understanding, expected to be signed later this summer, followed in the autumn by a detailed description of each country's contribution to the project. The entire package must be ratified by the board in November, she says, to keep the project on schedule.

One technical issue that must be resolved quickly is the final design for the two telescopes, which will make infrared as well as optical observations. Three companies are competing for the type of mirror to be used in the telescopes, two US companies and one from Germany.

The telescope on Mauna Kea, Hawaii, is expected to see first light by the end of 1998. Its nearly identical counterpart, adjacent to the Cerro Tololo Inter-American Observatory in Chile, should become operational two years later.

However, that schedule depends heavily on whether NSF can deliver its share of the funding on time. The project received only \$12 million of the \$16 million requested for the current fiscal year, which ends on 30 September, after Congress reduced NSF's overall budget request. The agency's re-

quest for next year, now before Congress, will almost certainly be trimmed again.

"Last year Gemini was taxed to support other large projects in the physical sciences", says Bless. "But NSF has assured us that they consider the project to be very important, and the fact that it's an international effort gives it a high visibility."

Wolff is not quite so optimistic. "The cuts won't hurt us this year", says Wolff, "but next year the impact would be felt on the telescope's enclosure and its primary mirrors. Even if we get our request for FY [fiscal year] 93, it's not clear that we will have enough money to do the job right. We really needed \$21 million to make up for what we lost in 1992, but we only asked for \$17 million."

NSF is the executive agency for the project, which is being managed by the same coalition of research universities that runs NOAO. But Julie Lutz, director of NSF's astronomy division, says that it is harder for the United States than for other countries to sustain a long-term scientific collaboration because the entire US budget is reviewed annually by Congress.

Wolff says that she may need more than six months to accomplish everything that must be done. At the same time, she says that scientists must know by the end of the year how much they have to spend. Any delays, of course, invariably increase the cost of the project.

Jeffrey Mervis

Star wars — the astronomers strike back

London. Astronomers are becoming more aggressive in trying to protect their 'laboratories' from pollution.

At the end of this month in Paris, researchers at a conference organized by the International Astronomical Union (IAU), the United Nations Educational, Scientific and Cultural Organization (UNESCO) and the International Council of Scientific Unions (ICSU) will try to extend legislation protecting unique Earth-bound sites to locations with especially good skies or low levels of radio noise. A week earlier, astronomers attending a symposium as part of the University of Chicago's centennial anniversary will propose a multilateral treaty to protect near-Earth space from the hazards of satellite debris, a measure that would considerably benefit astronomers.

The Paris conference is ostensibly an occasion to debate the issues of pollution on an international stage. It is expected to close with a declaration on the need to reduce adverse environmental impacts on astronomy. A pre-conference draft calls for reduced levels of outdoor lighting, a sus-

tained presence in the competition with the communications industry for radio wavelengths and measures to reduce and remove space debris.

One new tactic that astronomers have adopted is to seek an extension of the World Heritage Site programme, run by UNESCO and the International Union for Nature Conservation. At present, this programme is used to protect valued buildings or natural environments, but not such intangibles as light and radio characteristics.

Status as a World Heritage Site does not guarantee protection — Stonehenge in England, for instance, is protected but still suffers as a result of alterations made to improve tourist access. However, designation does provide a degree of international legal recognition and a basis from which to fight possible depredations.

Radioastronomy has been increasingly harassed by the communications industry with its demands for more and more radio frequencies on which to operate and its increasing use of satellites. In the past, observatories could be hidden from ground-

based sources by building them in valleys, but there is no escape from radio signals being beamed down from on high.

Despite facing great pressure from the communications industry, the International Telecommunications Union (ITU), the body responsible for regulating access to radio frequencies, reaffirmed earlier this year its commitment to keep clear certain frequencies used by radioastronomers. The ITU has also recognized for the first time the problems caused by badly filtered communications equipment.

A particularly severe source of radio noise are the GLONASS navigational satellites launched during the 1980s by the Soviet Union. These satellites were intended to create a global net from which pilots and mariners could triangulate their position at any time. However, it soon became clear to radioastronomers that these satellites were virtually unshielded — although transmitting in the 10–20 MHz range, the emissions could be detected up to 100 MHz away — and that each operated at a different frequency.

Ian Mundell

New foundation to fund Russian biologists

Moscow & Washington. An international foundation has been formed to keep intact an elite group of Russian biomedical researchers during the current economic crisis in their country. Acting on the assumption that there are no more than "20 to 30 basic biology labs" in Russia doing first-rate work, Alexander Goldfarb, a biochemist at the Public Health Research Institute (PHRI) in New York, is leading an effort to raise \$5 million over the next five years for peer-reviewed grants to teams of Russian scientists.

The idea for the American-Russian Biomedical Research Foundation (BRF) grew out of an existing scientific collaboration between Goldfarb, who left the Soviet Union for the United States in the 1970s, and Vadim Nikiforov of the Institute of Molecular Genetics (IMG) in Moscow. In April, the Soros Foundation, a private philanthropy created by George Soros, a wealthy US financier, contributed \$100,000 to cover BRF's start-up costs, while the Russian Academy of Sciences and the Ministry of Higher Education, Science and Technology have each provided 350,000 rubles (\$3,000). Soros has also given \$50,000, to be converted into rubles, to supplement the salaries of dozens of Russian scientists working collaboratively with US researchers.

Goldfarb, who is president of the new foundation, says that it will operate sepa-

rately from governments so that the money is not "mainly wasted on the machinery" of administering research. A panel of prominent US scientists will oversee the process of selecting the most promising projects from among proposals submitted to the IMG and then forwarded to the West.

PHRI, a nonprofit research organization conducting research in biology and medicine, will be the foundation's base in the United States. This year its two-person staff will concentrate on raising funds from companies and private organizations concerned about the future health of Russian biomedical research. The Moscow office at the IMG will meanwhile solicit grant applications. If its fund-raising efforts are successful, the foundation would like to make ten research grants, of \$50,000 each, by 1 September or soon afterwards.

Goldfarb believes that many potential contributors have been deterred because they did not know who to give their money to and whether the funds would be spent properly. BRF will provide the infrastructure to overcome those problems, he says. Each grant will carry a 40 per cent surcharge, says Goldfarb, "so that we can provide the infrastructure that doesn't exist in Russia". The money will pay for such administrative functions as an accounting and purchasing system, as well as the cost of delivering needed

equipment and supplies.

The first test of BRF's infrastructure is a plan to distribute \$1,000 to each of the 50 or so Russian laboratories collaborating with US partners. With the collapse of the economies of the countries in the former Soviet Union, the US National Science Foundation provided each US grant recipient with an additional \$10,000 to buy equipment and supplies for its Russian partners. The Fogarty International Center of the US National Institutes of Health sponsors a similar programme. But both agencies prohibit US scientists from paying any part of the salaries of their foreign collaborators.

One difficulty, not yet resolved, is that grant recipients could find their regular budgets reduced and the funds used to support other projects. Russian research managers have considerable discretion over the distribution of rubles, and researchers are worried that BRF could create a situation in which the stronger laboratories are financed by the West and the weaker ones supported by the Russian government."

Goldfarb plays down that possible division, saying that the foundation is only a short-term solution to the problems facing his former colleagues. "Once the Russian economy improves", he says, "then this type of outside support won't be necessary."

Vladimir Pokrowsky & Jeffrey Mervis

US realigns misconduct offices at NIH

Washington. The US government announced last week that it will create an Office of Research Integrity as a compromise solution to a continuing dispute over the operation of the scientific misconduct office at the National Institutes of Health (NIH).

The decision to close the Office of Scientific Integrity (OSI) and move its operations out of NIH and into its parent agency, the Department of Health and Human Services (HHS), is meant to address criticism that OSI, as a part of NIH, was too close to the scientific community to be impartial. It will also answer those who argued that OSI lacked the regulatory authority both to investigate and to adjudicate, and that it failed to provide the necessary legal protection to the targets of its investigations. The new Office of Research Integrity will combine the staff and functions of both the OSI and the Office of Scientific Integrity Review (OSIR), an oversight office that has reviewed OSI investigations and proposed sanctions in individual cases.

Those two duties will be better defined and separated under the new arrangement. The investigative division will do the work of the old OSI, but its reports will contain

only investigative results and recommended conclusions rather than reaching a conclusion about whether misconduct has occurred. If an investigation does result in a proposed finding of misconduct, the accused researcher can ask for a special hearing, similar to that offered to those who, for whatever reason, face the loss of federal funds.

The new hearings will offer scientists such due-process protection as access to evidence and to an attorney and the right to cross-examine witnesses and to present rebuttal evidence and witnesses. The hearing panel will consist of an HHS administrative law judge and, in cases where the judge wants expert advice on complex scientific issues, two independent researchers. More people will be working on investigating misconduct, especially on the legal side: the new investigative branch will employ six lawyers, compared with one on the current OSI staff.

J. Michael McGinnis, now director of the office of disease prevention and health promotion within the Public Health Service, has been named acting director of the office. A search is under way for a permanent

director. Lyle Bivens, the current director of OSIR, will direct the policy division of the new office. Although Jules Hallum, who now runs OSI, has been named acting director of the new investigative division, he may not remain. Hallum, a virologist, is 68 years old and is thinking of retirement.

Hallum fought to keep OSI at NIH, and he is concerned that the merger of OSI and OSIR may reduce the credibility of the investigative process. Although an independent hearing serves an important role when misconduct is uncovered, he says, he is unhappy that no independent review will be conducted when there is no finding of misconduct. Without OSIR as a double check, he asks, "who's going to prove that we aren't whitewashing people?"

Hallum also says that stripping OSI of its adjudicative aspects could "let us get away with less due process" in the investigative phase. But he promises that there will be no reduction in such protection while he is running the operation.

The new office, which is to open on 15 June, will be located in HHS's offices in Rockville, Maryland.

Christopher Anderson

Alternative toxicity tests

SIR — Your correspondents (*Nature* 341, 680; 1989 and 353, 489; 1991) have described plans for replacing the LD₅₀ toxicity test with the “fixed-dose procedure” first advocated by the British Toxicology Society, which avoids using death of the animals as an endpoint, and instead relies on the observation of clear signs of toxicity developed at one of a series of fixed dose levels. This welcome development, promising revisions of international guidelines, has taken a decade to reach fruition, partly because of the time needed for the worldwide evaluation prompted by the British society.

In Germany, there is parallel interest in a procedure, originally suggested by Professor Gerhardt Zbinden, that would make still more economical use of animals in toxicity testing. This method also uses a series of fixed dose levels, but makes use of a smaller group size of three per dose range and a sequential testing strategy that requires at each dose level a decision either to stop further testing or to proceed at a higher or lower dose level. Validation tests now under way suggest that the “acute toxic class” procedure requires an average of only 7.5 animals per test, compared with 17.5 per test in the fixed dose procedure — and many more in the LD₅₀ test.

Laboratories in the United States and Japan are collaborating in tests of the procedure. We hope that, in due course, EC and OECD guidelines will be further revised so as to accommodate it.

Horst Spielmann

*Institut für Veterinärmedizin
des Bundesgesundheitsamtes,
Postfach 33 00 13,
D-1000 Berlin 33, Germany*

Hungary's plans

SIR — Your recent articles on Hungarian science (*Nature* 355, 669–672; 1992) give a one-sided picture. This may be due to simplification, preconceptions or a lack of information. After almost 50 years of military and political confrontation, it takes time to change prejudices both in the West and in the East. So let me draw the other side of the picture.

A recent study by the Organisation for Economic Co-operation and Development (OECD) noted that “there is a strong tradition of respect and support for knowledge and science in Hungary”. We have a great many qualified people, who are well respected at international conferences and research laboratories. Following the collapse of the Soviet system, the state of technology transfer

is relatively poor, but the OECD study applauds the plans of universities to set up “Fraunhofer-like” institutions and technology centres close to their campuses. OECD also remarks on Hungary's efforts to change its international orientation. In 1991 alone, the government signed an agreement on cooperation with the European Space Agency, became a full member of COST, started project-level participation in some EUREKA projects and successfully used the PHARE assistance programme for improving its scientific infrastructure. The OECD study noted that “this cooperation constitutes the most efficient vehicle for integrating Hungarian researchers in international networks”.

Several years ago, Hungary set up new competitive funds for financing research and development, “important tools for stimulating and guiding the research effort”, the OECD study says. Moreover, the academy, which has played a very important role even in the past 40 years, continues to work for the improvement of Hungarian science. According to the OECD study, “recent and planned . . . reforms make it fully independent from the government and enlarge its scientific constituencies . . . Significant actions recently taken or being planned are going in the right direction.”

Lajos Nyiri

*National Committee for Technological
Development,
H-1052 Budapest,
V. Martinelli ter 8, Hungary*

Fetal tissue

SIR — Leonard Hayflick (*Nature* 356, 652, 1992) points out that many of the people who support the current US ban on the use of federal funds for transplantation of aborted fetal tissue into patients may be benefactors of research that he and others have conducted using normal human fetal cell strains. This begs the more relevant question, which is whether or not such people would have supported this research if they knew that it depended on the intentional destruction of the ‘donors’. Hayflick does not clarify whether his own work, or earlier work, relied on the use of intentionally aborted fetal tissue or whether the same work was, or could have been, carried out with tissue obtained from spontaneous abortions.

The failure to make such a distinction suggests that he misunderstands the objections raised by those of us who support the administration's ban. We believe that the means of death of the individuals from whom the organs, tissue or cells are obtained is morally relevant, regardless of the age of the ‘donor’. We

object to the use of tissue derived from induced abortion because of the way in which it is obtained.

The most effective argument against the notion that we can somehow separate the practice of abortion from the utilization of its ‘products’ is presented by Hayflick who suggests benefactors are in no position to raise moral objections. In other words, there is a link between the acquisition and subsequent use of the tissue. But the existence of benefit does not undermine the basis for objecting to the manner in which the research was conducted, and if informed that the benefit depends on an immoral practice, many of us would forgo the benefit.

However, Hayflick raises an interesting point. As a society, it is not clear how we can raise any meaningful objections to the exploitation of fetal life for any purpose (commercial or other) as long as we maintain the policy that such individuals are expendable. We find ourselves in the ironic position of valuing the human fetus as a contributor to society but not as a member with protectable interests.

Keith A. Crutcher

Department of Neurosurgery,

William Everson

*Department of Obstetrics and Gynecology,
University of Cincinnati Medical Center,
Cincinnati, Ohio 45267–0515, USA*

No politics please

SIR — I have long considered *Nature* to be the preeminent source of news on contemporary scientific issues. Feeling the need for neoconservative doctrine, I would consult a journal, such as *The Economist*, specializing in such material. I was therefore dismayed to read in the Opinion section of *Nature* (356, 92; 1992) a political commentary calling for an immediate free trade agreement between Canada and Mexico. The notion that such an agreement would be beneficial for Canada ignores the massive de-industrialization experienced by Canada following the recent agreement with the United States. Furthermore, the writer seems to be unaware that Mexico is governed by a dictator. An unqualified belief in the benefits of free trade, irrespective of the political legitimacy of the Mexican government, betrays a faith unlikely to be shaken by unpleasant facts. A scientific journal should not be a forum for the sort of unsustainable assertions which seem to gain credibility through repetition: acclamation is not proof of worth.

Paul Cumming

*Montreal Neurological Institute
and Hospital,
3801 rue University,
Montreal, Quebec, Canada H3A 2B4*

NSF stance on 'skills shortage'

SIR — Jeffrey Mervis's article on the controversy surrounding trends in degree acquisition (*Nature* 356, 553; 1992) deserves comment.

Having come to this issue after the fact, I have had the benefit of examining it with a fresh perspective and have approached it in as objective a manner as possible. Nothing in my review suggests to me that there was a deliberate attempt on the part of the National Science Foundation (NSF) to misinform.

Both our friends and critics ask how this narrow study of degree attainment could assume the status of an authoritative labour market projection of demand for scientists and engineers. In response, I would say that, with the benefit of hindsight, the shortcomings of this limited analysis seem self-evident. The widespread use of the study beyond that which its narrow scope could support has served for me as an object lesson that sins of omission can be as important as sins of commission. The NSF now has in place procedures to forestall a similar occurrence. Publications released under the NSF imprimatur will meet the high standards we have set for ourselves and the expectations others have of us.

NSF continues to give high priority to improving the quality of education. The contention that we can have too many citizens with a good grounding, at various levels, in mathematics, the sciences and engineering is untenable, given the broad challenges of our developing economy in the years to come. With this in mind, we will continue to conceive and support programmes to nurture technical knowledge and skills and ensure the development of the scientific and engineering talent to meet future needs.

Walter E. Massey

(Director)

National Science Foundation,
1800 Q Street, NW,
Washington, DC 20550, USA

Lithuania's need

SIR — I write in response to your leading articles of 23 January and 20 February (*Nature* 355, 283 & 659; 1992). The latter referred to the "disconcerting . . . tendency for the world to be sprinkled with small countries" and seemed to support those who would argue that small countries should be ineligible by reason of size to self-determination and, as victims of oppression, are untrustworthy to govern themselves and their minorities fairly.

You assert that minorities in the newly

independent "microstates" are in immediate danger of oppression and specifically question whether the Baltic states are protecting the rights of their Russian minorities. Our law clearly states that Lithuania "shall guarantee to all its citizens, regardless of ethnicity, equal political, economic and social rights and freedoms, shall recognize its citizens' ethnic identity, the continuity of their culture, and shall promote ethnic consciousness and the expression thereof."

These expressions include state-supported Russian, Polish and Hebrew language schools, contact with people of the same ethnic background abroad, and religious or folk observances in one's native language. In short, there is not, nor has there been, any oppression of national minorities in Lithuania.

Furthermore, contrary to the reasoning in your leading article of 30 January (*Nature* 355, 377; 1992), the scientific and economic value of the former Soviet Union is immeasurably exceeded by the sum of its smaller parts. In order to be able to sustain itself in the short term and to prosper in the coming years, Lithuanian science needs the journals, the travel, the research support now being advocated only for Russia. A modest sum of money concentrated in a small country goes much further than a large amount of money scattered throughout a massive country.

Stasys Lozoraitis

(Ambassador to the

United States of America)

Lithuanian Embassy,
Washington, DC 20009, USA

Life on a grant

SIR — Howard Morris (*Nature* 356, 10; 1992) and M. J. Clemens (356, 280; 1992) referred to the increasing difficulty of recruiting research students without specifically identifying the biggest disincentive to postgraduate study, namely the ludicrously inadequate size of maintenance grants. At the launch of *Nature's* "Manifesto for British Science" last September, the head of the Science and Engineering Research Council (SERC), Sir Mark Richmond, declared that the annual stipend for PhD research students was "a national disgrace".

Since his comments, both SERC and the Natural Environment Research Council (NERC) have announced small rises in their basic awards to £4,300 a year, with NERC adding a further increment next October of £500 to first and second years and £1,000 to third years.

Thus, a maintenance grant from NERC or SERC (taking income tax, national insurance and 'poll tax' allowances into account) is equivalent to full-time employment on less than £5,500 per

year. Compare with the average starting salary of £11,681 for a graduate in science and engineering or the £7,000 that would be earned by the lowest paid full-time workers if there was a national minimum wage of £3.40 an hour.

Is it any wonder that, after striving for three or four years to attain a first or upper second class degree, graduates decide against another three years of debt accumulation to obtain a PhD? Indeed, it is amazing that anybody can afford to become a postgraduate after experiencing the abject poverty that accompanies undergraduate life these days. It may well be difficult to convince graduates that they should forsake a living wage in favour of academic research at the moment, but it will surely be an even more arduous task once the graduate job market recovers from its present depression.

However, just as Morris claims that the Committee of Vice-Chancellors and Principals has been reluctant to "argue the case vigorously for fair rewards for Britain's academic teachers and researchers", postgraduates may make the same accusation of their paymasters, the research councils. What are we to make of the current lack of agreement on the rate of awards for three-year studentships? SERC offers £4,300 a year, NERC £4,800, the Agricultural and Food Research Council £5,000 (soon to be £5,500 or more), and the Medical Research Council £5,590.

I agree with the recommendation in *Nature's* manifesto that a "doubling of present rates would be simple and equitable". Who else shares this view?

Bob Ward

Department of Geology, The University,
Manchester M13 9PL, UK

Born to it

SIR — Richard Peto's Commentary (*Nature* 356, 557; 1992) misses an important aspect of how to go about controlling chronic diseases that kill adults in middle age. David Barker at Southampton University in the United Kingdom has discovered a relationship between various perinatal factors and the subsequent likelihood of developing certain chronic conditions in adult life. Placental mass, for example, seems to be independently correlated with adult blood pressure decades later, and the duration of breast feeding seems to be correlated with adult blood cholesterol levels. Perhaps, to follow the course Peto suggests, we need to pay more attention to the conditions in which children are born.

Olivia Judson

New College,
Oxford OX1 3BN, UK

Reflections of a whistle-blower

E. J. R. Rossiter

Scientific fraud, much in the news recently, is not a new phenomenon. What follows is a personal account of a particular case.

SCIENTIFIC research is a highly competitive process, and, unfortunately, not all research is honestly carried out. This is the story of a man of forceful personality who moved from one academic job to another. It is a story of failure to check on his previous appointments and failure honestly and thoroughly to assess the merit of his scientific work. In the light of the scandal that is now emerging surrounding William McBride, a Sydney gynaecologist, I wish to draw attention to the problems that, in my own personal experience, can be associated with 'whistle-blowing'.

Michael Briggs, an English academic, was appointed foundation professor of biology at the newly formed Deakin University at Geelong in Australia in 1977, the year in which it first took students. I became a member of the university council in 1978, remaining until 1987 when my term expired. In 1979 I joined the newly formed ethics committee, becoming its chairman in 1982, a position I still hold. In mid-1979, I came across a list of Briggs's publications. I was struck by two things. Many of the papers appeared to be review or invited articles, while others were single-page references suggesting that they were items of correspondence. Later, when I became concerned about Briggs, I obtained copies of all the publications listed in his *curriculum vitae*. Many of them appeared in proceedings of meetings sponsored by drug companies, and several journals in which he had published his research were owned by such companies. Only three papers might have been considered original research and were published in refereed journals. Those that appeared in internationally respected and refereed journals such as *Nature*, *Science*, *New England Journal of Medicine* and *The Lancet* all turned out to be ephemeral contributions.

During 1982 and 1983, several people made comments to me about Briggs. At the same time I became aware that he was frequently away from the university and, it turned out, almost never spent a full day there. Whereas I realized that many academics spend a considerable amount of time away, it did concern me that a professor and foundation dean of a department should spend such a small amount of time doing the work for which he was paid.

It has been said that Briggs comman-

ded international respect for his work and was highly qualified. Apart from his previous professional appointments, his two senior degrees were PhD (Cornell) and DSc (New Zealand). A PhD is a senior qualification, routinely obtained from research at a recognized institution over a three- to four-year period. A DSc, however, is usually given to a scientist of distinction on the basis of research, published papers and standing in the scientific community. On writing to the universities concerned, I discovered that Briggs spent only about a year at Cornell as a tutor and received an MSc, not a PhD. He then spent less than three years at what was then the University of New Zealand. He did in fact receive a DSc while there: it is unusual for this degree to be awarded in such a short period of time, and one wonders how much this award was dependent upon Briggs's PhD.

A side issue, which is nonetheless important, is the failure of the selection committee at Deakin to check Briggs's qualifications before his appointment. Failure to check qualifications has been noted in other appointments at reputable universities around the world where workers were later to be found conducting fraudulent research.

My first official act in this matter was on 28 October 1983 when I wrote to the chancellor of Deakin University, Justice Austin Asche. I brought to his attention a confidential file of papers, including comments from the then president of the Royal Australasian College of Physicians. I expressed my concern that unless the university set its house in order, a journalist might discover what was happening and the university might be damaged. I pointed out that I could find no respected specialist in Briggs's field who spoke well of him, and also that I considered he could easily sue the university. I suggested various eminent, independent people who could be consulted for advice.

The chancellor invited me to his chambers in Melbourne. I found him disbelieving, and he declined to consult any expert on a confidential basis. Following this discussion, and on the advice of Asche, I wrote a courteous letter to Briggs on 28 November 1983 in my capacity as chairman of the university ethics committee, expressing my concern and offering to speak to him personally

before discussing the situation with the committee. I pointed out my authority under article 3 of the ethics committee to take this step. I also made him aware that I had heard unfavourable comments from inside and outside the university and the country, and set out a list of seven questions. I received no reply.

On 12 December 1983, I recommended to the chancellor that a committee of inquiry be set up. He disagreed, stating that: "I cannot too strongly emphasise that nothing is so far proven against Professor Briggs and it is elementary that a man is innocent until proved guilty. It seems to me that every opportunity should be given to him to answer the enquiries you have raised before any further steps are considered."

The reply from Briggs to my second request is of interest and for that reason I include it here in full.

Thank you for your letter of 28 November. I am sorry for the delay in replying, but I have been overseas.

I am aware that I have been seriously slandered by a member of Deakin staff, both locally and nationally. I am seeking legal advice on action for defamation.

I am most happy to supply the information requested in your letter. The paper to which the information requested in your letter refers was one of a series investigating dose-dependent biochemical effects of contraceptive steroids. It was prompted by the announcement in 1970 of the Committee on Safety of Medicines of the parent relationship between oestrogen dose and risk of thrombosis.

Preliminary studies were undertaken at the University of Sussex and the Ulster Polytechnic. The protocol, however, was not finalised until 1971 and recruits were then sought. The study was initiated at the University of Zambia, Lusaka, and women were recruited through the staff clinic and through a large general practice in the city. The study was continued at the Alfred Hospital, again with recruitment largely through the staff clinic, until 1975.

A copy of the informed consent form is enclosed. This was approved by WHO. At each of the two visits, volunteers received payment for travel expenses. Approximately 8 ml. of blood was collected on each occasion. This was divided into a series of aliquots in the laboratory and what was not used immediately was preserved by deep freezing.

The study was funded by grants from the University of Zambia, WHO, IAEA, Schering AG and Wyeth International.

The proteins listed in the paper were

measured in the Biochemistry laboratories of the University of Zambia and the Alfred Hospital, with some exceptions. Assays for transcortin, SHBG, TBG and RBP were not available to Zambia and were all carried out at the Alfred. Unfortunately the results for TBG were unreliable and eventually had to be repeated by a contract laboratory in the USA (ISI, Los Angeles).

Final compilation and statistical analysis of the results was carried out in 1978 and the results presented at the Fifth International Congress on Hormonal Steroids held in New Delhi in November 1978.

The paper was initially published in the proceedings of this meeting (edited by V.H.T. James and J.R. Pasqualini), which constituted a special volume of the *Journal of Steroid Biochemistry*. Instructions to authors asked for the current correspondence addresses of all authors, which is why the Geelong Hospital is mentioned.

I believe that the above account provides all the information you requested, but I would be most happy to supply anything further should some aspects of the research still be unclear.

May I say that I welcome surveillance by the Ethics Committee and trust that my reply is satisfactory.

Following this I made the formal complaint to the vice chancellor of the university which is set out below.

Dear Professor Jevons

I am writing to you in my capacity as Chairman of the Deakin University Ethics Committee.

I would draw you to the Ethics Committee Terms of Reference Article II, paragraph B, where it is stated they "shall ensure that any research or investigation conforms to generally accepted moral and scientific principles".

And to Article III. The Committee may exercise such surveillance of teaching investigations, or research projects as it deems necessary, in order to ensure that the above requirements are being met. Such surveillance may be carried out by periodic review, spot checks, or otherwise, as the Committee thinks fit.

In my capacity as Chairman I have been approached by persons within and without Deakin University with certain allegations regarding Professor Michael Briggs. In view of the seriousness and delicate nature of the complaints I chose to make discreet inquiries without involving the Committee in the discussion of a named person.

Following this I wish to make a formal complaint and suggest an inquiry as set out under Deakin University regulation 3.3 (1).

I would refer to the paper in the *Journal of Steroid Biochemistry* Volume 2 pages 425-428, 1979 on Oral Contraceptives and Plasma Protein Metabolism. I have contacted Professor Briggs by letter and received a reply from him on 20 December 1983. I find it difficult to accept the reply which he has made.

My first concern is that I have difficulty in believing that eighty-eight women could be recruited through the University of Zambia in Lusaka through the Staff Clinic in a large

general practice and Professor Briggs states that the Study was continued at the Alfred Hospital through the Staff Clinic.

My second concern is the difficulty in carrying out the assays mentioned in this Study and these assays are as follows:-

Which Laboratory carried out the following analyses?

Transferrin
Ceruloplasmin
Haptoglobin
Transcortin
Sex Hormone Binding Globulins
Thyroxine Binding Globulins
Renin Substrate
Fibrinogen
Coagulation Factor 7
Coagulation Factor 8
Antithrombin 3
Plasminogen
Pre-Albumin
Albumin
Retinol Binding Protein
Lipo Protein Fractions

I have reason to believe that some of these assays could not have been carried out at the time of this Study in the way which has been suggested and Professor Briggs in his reply has certainly not made it clear to my satisfaction.

The second Study which I would raise is entitled a Randomised Study of Metabolic Effects of Four Oral Contraceptive Preparation containing Levo-norgestrel plus Ethinyloestradiol in Different Regimens. This is published in the proceedings of a special symposium held at the tenth World Congress on fertility and sterility Madrid July 1980 by M.T.P. Press Limited edited by R. B. Greenblatt.

Again it is extremely difficult to understand how 80 women were recruited for this study in view of the very considerable difficulty experienced by most research workers in recruiting adequate numbers of subjects for this type of work and I believe the enquiry should seek to know precisely how these women were recruited. I am also concerned about the medical participation in a study of this size and would like confirmation that one medical practitioner was involved in all of this work.

I would like to know where the measurements of the hormones LH and FSH (Table 6.2) were carried out and the date at which they were done.

In the measurement of blood coagulation factors particularly fibrinogen, Factor VII, Factor VIII, Factor X and Antithrombin III and Plasminogen also in Chapter 6.2 are difficult estimations to be done by any laboratory and extremely demanding.

I believe the question should be asked and proven as to which laboratory carried out these investigations.

Apart from the factual information which I believe is required and which has not been forthcoming from Professor Briggs, I think there are serious problems concerned with the research interpretation and I do not believe that any member of the Deakin University staff or the Ethics Committee staff is competent to consider all of the issues in this very specialised field and my list of recommended persons is as follows:

Apart from the six people I recommended, I suggested to the committee of

inquiry that a list of specialists should be drawn up who could be consulted.

The complaint was duly taken to a committee of inquiry, properly constituted, where Briggs was supported by his union, the Federation of Australian University Staff Association. The university visitor, a mediaeval anachronism in the University Act, was asked to intervene. The visitor in fact was the Governor of Victoria, and he appointed Mr Justice Stark of the high court to give an opinion, which was, in brief, "that the complaint was improperly founded". Following this, a further complaint was drawn up with the assistance of Professor Bryan Hudson and Professor Henry Burger and a second committee of inquiry was created. Just before that committee was due to meet, Briggs resigned from his university appointment and, under the University Act, no further inquiry could legally be held by the university. He disappeared overseas and died in 1986 in the Costa del Sol, Spain, of liver failure. Nevertheless, the university did set up a committee in 1984, chaired by a member of the university council, Marjorie Ramsay, to look into the case. The committee ultimately produced a long series of recommendations about the university's inquiry procedures.

I would like to express some of my concerns for what happened in the next few months. During the period of the first inquiry and of Briggs' defence I became subjected to a barrage of attacks and criticism. I soon learned that I was a "troublemaker" and "a stirrer" in attacking one of the outstanding academics of the university. People on the council ignored me when I spoke to them, particularly members of the academic staff. At no time was the university council aware of the substance of the allegations made to the committee of inquiry, which were confidential.

In 1984, at a meeting of the council of the university, the composition of council committees was discussed. My re-nomination as a council representative on the tenure committee was opposed by Dr Michael Stokie, dean of science, on the grounds that I would not be an unbiased member. I believe that opposition to an appointment as a representative of a school on these grounds was unprecedented. Rather than be involved in further conflict, I withdrew.

On 26 June 1985, Professors Henry Burger, Bryan Hudson and I lodged a second complaint with the university which I set out below:

We hereby lodge a formal complaint alleging that a member of the staff of the University, namely Professor Michael Briggs, has been guilty of gross misconduct.

(1) In publication Briggs & Briggs 1980

Reproduction Vol. 4 (Suppl) 79 and Briggs Michael "Bulletin of the Post Graduate Committee of Medicine" University of Sydney 1980, 36, 148 Professor Briggs alleged that a compound known as desogestrel was tested and compared with another compound. At the time of the alleged study desogestrel was not available in Australia, and Professor Briggs did not have the necessary import licence to import the compound. We allege that Professor Briggs did not carry out the said tests.

(2) In late 1983 Professor Briggs advised Organon (Australia) Pty. Ltd. the manufacturers of desogestrel that he had done research into, *inter alia*, desogestrel. He offered to make the results available to the manufacturer in return for a specified grant of money. A sum of money was paid by the manufacturer to Professor Briggs' private account, and not as should properly have been done, to the funds of Deakin University. (3) Professor Briggs published papers in the *Bulletin of the Post Graduate Committee on Medicine of the University of Sydney* 36: 148, 1980 and in "New Considerations in Oral Contraception" Ed: 1 Brossens P9, 1981 and in conjunction with Dr Maxine Briggs entitled "Contraception 23" 463 1981.

In the said papers Professor Briggs claimed that a total of 291 volunteers, who had to fulfil strict criteria, participated in the studies which were the subject of the papers. No such number of volunteers did in fact exist and we allege that the said studies were not carried out to the extent claimed, if at all.

(4) No approval was given by the University Ethics Committee for the research projects referred to in the preceding paragraph.

(5) In the first publication referred to in paragraph (3) hereof, Professor Briggs refers to measurements of blood constituents and has claimed that these were carried out at the Alfred, Geelong and Prince Henry's Hospitals. No such measurements were carried out for Professor Briggs at any of the said Hospitals and we allege were not carried out at all.

(6) In conjunction with Dr Maxine Briggs, Professor Briggs published a paper in the Proceedings of the 10th World Congress on Fertility and Sterility Symposium entitled "The Development of a new Triphasic Oral Contraceptive P. 79 1980".

In the said paper Professor Briggs claimed that measurements of follicle stimulating and luteinizing hormones were undertaken by Dr K. Fotherby. The said measurements were not carried out by Dr Fotherby, and we allege were not carried out at all.

(7) In a publication in *Res. Veterinary Sciences* 28: 199, 1980 Professor Briggs claimed that in the research the subject of the said publication, beagle dogs were used to measure progesterone receptors. By letter dated 10/3/83 Professor Briggs claimed that the measurements of progesterone receptors were carried out at The Huntingdon Laboratories in the United Kingdom. No such measurements were taken either at the said Laboratories or at the University and we allege that no such measurements were made at all.

It is submitted that each of the foregoing constitutes gross misconduct on the part of Professor Briggs and we request that you instigate the prescribed procedures for the investigation of such matters.

Professor's suspect investigations supported the Pill's safety claims

WHEN Professor Michael Briggs arrived at Deakin University in 1976, his background was beyond reproach.

The British scientist had a string of publications on the contraceptive pill. To Prof. Briggs the then vice-chancellor, he was the sort of academic who might put the university in the map.

It was Briggs, who really started to get research going and who built up the school of sciences into what it is today.

Prof. Briggs, who brought funds from drug companies to buy equipment for students to work with personally, he was charming and we often had dinner together.

He was now aged 51, unmarried.

Pill also affects menstruation, setting a regular cycle through hormone control, varying the chemical content of the pill can also cause erratic, break-through bleeding.

The hopes of many researchers for further reducing the risk of vascular disease were raised with the discovery by the West German pharmaceutical giant Schering AG of a new way of producing pills, delivering three varying doses of hormones over the menstrual cycle.

Although other products make competing claims, these pills, known as "triphasic" and containing the synthetic progesterone (norgestrel), have been extensively promoted by drug companies as an even safer form of birth control.

The Briggs case

After one of the biggest academic scandals ever in Australia, Professor Michael Briggs resigned from Deakin University last October and disappeared overseas, preventing an inquiry by the university into allegations that he fabricated research on the contraceptive pill. Now the *Sunday Times* has tracked him down to the Costa del Sol, and he refused to be interviewed.



proved against Professor Briggs, declared Mr Justice Keele, the chancellor, accepting the resignation. "Nothing adverse to his reputation has been established and no inference adverse to him should be drawn."

This was not, however, the verdict of the international scientific community. Last October, a special meeting was held in West Berlin, solely to discuss Briggs. British doctors agreed that researchers should no longer use his findings. They believe that the drug was discovered in his work are so serious that earlier papers relating to many more contraceptive products must now be re-examined.

"What is clear is that we must now trust any of the work he did," said the doctor.

Fraudulent work exposed by the *Sunday Times*.

During the period from the first complaint until the final resolution, I received about 200 telephone calls. On two nights I had telephone calls at 1.00, 2.00, 3.00, 4.00 and 5.00 a.m. Some of them consisted of heavy breathing, some were obscene, two threatened my life but only a few mentioned Briggs. After his death they stopped. Also during this time I received two letters on Deakin University stationery containing obscene allegations about my private life.

A further disturbing comment was made to me by two members of the council that a senior member of the academic staff, himself a psychologist, had, in one of the committees considering the affair, stated that in "his professional opinion Jim Rossiter was of unsound mind". This person never spoke to me about the Briggs affair nor discussed the reasons for my complaint. Also at this time, my private referral work diminished dramatically.

Peace might then have descended but for an article appearing in September 1986 in the *Sunday Times*, which had tracked Briggs down to Spain and in which he admitted generalizing small-scale findings to large, apparently convincing trials. There has been considerable speculation about how the newspaper became aware of the situation and from whence its information came. On two occasions I had come into my consulting rooms in the middle of the night and both times had found the door unlocked. I believe that the entire file on the Briggs affair was removed from my consulting rooms and photocopied by someone who was determined to expose the affair. Without condoning this illegal act, it certainly made my situation much easier. Immediately following the *Sunday Times* article I was interviewed on the radio. I was told by the then vice chancellor, Professor Skilbeck, that I could speak only as a private individual and not as a member of the university.

The fact is that a fraud of this nature is of enormous importance to the community. I had been scrupulous in trying to protect the university, while at the same time dealing with a man who had perpe-

trated criminal acts. However, I believe I conducted subsequent interviews with dignity and honesty even without support from the university.

A short while after this, the composition of committees of the university council for the following year was considered. It was not until the agenda arrived that I discovered that the members to compose the ethics committee had been completely changed without any consultation. At the council meeting I pointed this out and renominated all members of the previous ethics committee who wished to continue serving and who had been removed without consultation. All were re-elected. I was then myself renominated and re-elected unopposed by council. It is disturbing that some people would see it as important to remove not only those involved in the scandal but those who might have been involved in bringing it to the attention of the authorities.

Why did Briggs carry out his fraud? Was it for financial reward or personal glory? It does seem significant that all his research was related to products from two international drug companies. If a drug company sponsors research, what sort of supervision and monitoring should it and does it carry out?

There are many things that must remain unsaid. The laws of defamation preclude me from stating some of the most distressing aspects of this affair as is also the case in other major examples of scientific fraud. I am struck by the inability of many academics to accept the possibility that a senior person is dishonest; by the attacks upon personal integrity in a situation where the recipient is unable to reply; and by the behaviour of people commanding respect within the community who say one thing in private and another in public. The price of exposing fraud and deceit is high; it is sad, but necessary, that all stages of research must be guarded by vigilant people who must act without fear of the consequences of what they do. □

E. J. R. Rossiter is at 367 Ryrie Street, Geelong, Victoria 3220, Australia.

Reopening the solar neutrino question

A surprising new measurement confirms that the Sun shines but, confusingly, supports all plausible solutions of the long-standing solar neutrino problem.

THE solar neutrino problem has been with us now for 20 years, since Ray Davis and his colleagues announced that they had observed, at their underground Homestake detector, fewer than half as many neutrinos from the Sun as theorists had expected. The debate on the significance of that result has since engaged communities as different as stellar modellers and particle physicists, while the ingenuity of experimentalists has been tested by the development of different types of detectors. The latest measurements, announced just before this week's International Neutrino Conference in Granada, Spain, have caused a mixture of surprise, elation, consternation and confusion.

There are several sources of neutrinos in the Sun, of which the one directly linked with the solar luminosity is the nuclear reaction in proton-proton (p-p) collisions, which yields low-energy neutrinos. The other sources, due to other nuclear reactions (which give neutrinos of higher energy), can vary dramatically with the detailed astrophysical assumptions while remaining consistent with the observed luminosity.

For 20 years, solar modellers have argued that all the data on the Sun preclude a deficit of high-energy neutrinos large enough to explain Davis's result, so that the explanation must lie in some property of neutrinos. The sceptics, on the other hand, have argued that seeing is believing and that until the contradiction is shown to lie with a component of the neutrino flux tied directly to solar luminosity, the astrophysical estimates must be suspect. Both sides nevertheless agree on the importance of measurements sensitive to low-energy neutrinos, and since 1990 two international collaborations have been aiming to detect low-energy neutrinos from the Sun by their occasional conversion of gallium nuclei (specifically, the ^{71}Ga isotope) to germanium-71. The energy threshold for this reaction is 233 keV, so that it should be sensitive to a sizeable fraction of the p-p flux, which varies from zero to 420 keV.

During the mounting of these collaborations, other data on the solar neutrino problem have come from the Japanese Kamiokande detector, which detects bursts of Cerenkov radiation in a large underground water tank. Because Kamiokande detects neutrinos from the decay of solar boron-8 only, which is especially sensitive to the details of solar astrophysics, it could have been expected that Kamiokande would find a larger deficit than Homestake. But the opposite proved to be the case.

That has persuaded many to favour a neutrino-based explanation of the discrepancy — the so-called MSW (for "Mikhaev-Smirnov-Wolfenstein") solution. In this, a very small, but nonzero, neutrino mass causes neutrinos to 'oscillate' within the Sun between the three known species of neutrinos — the familiar electron neutrinos and those associated with the electron-like particles known as muons and taus. Unlike Homestake and the new gallium-based detectors, Kamiokande is weakly sensitive to muon and tau neutrinos, so that the higher flux it sees is understandable — at the cost of assigning mass to neutrinos, which implies dramatic new physics.

The first new measurements, which appeared last year, seemed to suggest that physicists should bite that bullet. The Soviet-American SAGE collaboration, at the Baksan neutrino observatory in the north Caucasus, uses 30 tonnes of liquid metallic gallium to probe for the p-p reaction. Six months of data yielded the striking result that the observed neutrino rate is only a small fraction of the expected rate and even, within the experimental error, is consistent with zero. The standard measure is the Solar Neutrino Unit (SNU), a likelihood of 10^{-36} per second that a target atom will interact with a neutrino. The standard model of the Sun predicts a measurement in the range of 125–132 SNU. Last year's SAGE result was $20^{+15}_{-20} \pm 32$, where the quoted uncertainties are statistical and systematic respectively. Only new neutrino physics could allow such a result.

Caution nevertheless suggested waiting for the outcome of the second gallium experiment, the multinational GALLEX experiment mounted in the Gran Sasso automobile tunnel, using 30 tonnes of gallium chloride (GaCl_3) in solution. On 2 June, the collaboration released two papers (preprints GX 1-1992 and 2-1992) submitted for publication to *Physics Letters*, announcing their preliminary results.

GALLEX is an elaborate device, as required of an instrument from which great sensitivity is required. The GaCl_3 solution is contained in one of two teflon-coated tanks for the length of a run, typically three weeks. Ga atoms converted to ^{71}Ge would reside in solution as GeCl_4 , which is more volatile than GaCl_3 and which is swept out of solution at the end of a run by nitrogen gas, converted to germane (GeH_4) and introduced into special miniaturized low-background counters.

The efficiency of Ge recovery, exceed-

ing 99 per cent, was checked with stable germanium added to the solution and with unstable isotopes produced by radioactive sources *in situ*. A frustrating background of cosmogenically produced, long-lived ^{68}Ge was effectively eliminated by no fewer than 28 successive extractions. Calibration with a nearby neutrino source is planned for next year.

On the basis of 14 runs (six of which have been fully monitored), GALLEX reports a neutrino flux of $83 \pm 19 \pm 8$ SNU (where the estimated uncertainties are statistical and systematic respectively). This is plausibly consistent with almost every candidate explanation of the neutrino discrepancy. The astrophysicists' estimate of the flux due to p-p reactions alone is 74 SNU, while the "minimal" solar models in which only p-p reactions generate solar luminosity predict 79 SNU. But within two standard deviations (2σ), the result is also compatible with the standard solar model. It differs from the median SAGE result by more than 2σ , but is consistent at 1σ with MSW predictions for a small region of the parameter space allowed by other experiments.

And the implications? Many physicists' quick supposition that, because GALLEX is consistent with the predicted p-p flux, astrophysics is to blame, is probably too quick. Even hypothetical models in which high-energy neutrino fluxes are arbitrarily reduced still fall well outside the 2σ limits of Homestake and the gallium experiments. For nonstandard solar models to be viable, at least one other experiment must be giving the wrong result. The MSW predictions are consistent at 1σ with all the experiments, but for a surprising corner of parameter space. And there remains the discrepancy between SAGE and GALLEX; until that is resolved, no firm conclusions can be based on either result.

That is why, among those assembled at Granada, those taking the most solace from the GALLEX data must have been experimenters working on the next generation of detectors, the Sudbury Neutrino Detector based on heavy water, for example. Meanwhile, the GALLEX result provides an unambiguous proof that p-p reactions (among others) power the Sun and that the solar neutrino problem is as slippery as ever.

Lawrence M. Krauss

Lawrence M. Krauss is at the Center for Theoretical Physics, Yale University, New Haven, Connecticut 06511, USA.

Practical Schrödinger's cats

Peter Knight

SCHRÖDINGER'S cat is alive and well. It is to be found living in cavities enclosed in superconducting metal, according to new work by Brune and colleagues in *Physical Review*¹.

Schrödinger described his cat to illustrate a paradox in quantum mechanics: a real cat can be alive or dead, but a quantum cat, hidden from view in a box, and which may have been poisoned, remains in a state of suspended semi-animation, neither alive nor dead, until someone lifts the lid to have a look. Brune *et al.* show how a more practical cat, a microwave field, can be made to exist in two distinct states, and demonstrate that it is possible to check that it remains that way.

Their experiment goes back to the problem of counting photons. Normally, detecting a photon simultaneously destroys it. But Braginsky and Khalili² and Caves *et al.*³ developed the idea of quantum nondemolition (QND) experiments in which photons could be sensed without destruction. The approach depends on the avoidance of back-action following measurement. The disruptive effect of back action is seen in attempts to measure the position of a particle: the measurement alters in an unknown way the particle's momentum, so that its subsequent path is unknowable. The trick in QND is to make a measurement in such a way that the change in the

measured quantity's conjugate variable (momentum in the case of position, the phase of electromagnetic radiation in the case of photon number) will not alter the quantity's subsequent evolution.

Developments in quantum optics have allowed these general ideas to be tested. One nondestructive method of counting photons is to look for a phase shift or energy shift, which depends on number and can be measured, without any absorption, in an interferometer. For example, the refractive index of certain media can be changed by a strong optical field. The resulting shift in phase of a second probe laser beam passing through the medium, measured interferometrically, can be used to count the photons in the stronger field⁴. This forms the basis of low-noise optical 'taps' used in optical communications.

Brune and coworkers adopt a similar approach to detect the very small numbers of photons in the microwave radiation stored in a superconducting cavity. Instead of a probe laser, they suggest using a beam of highly excited 'Rydberg' states of alkali atoms to sense and manipulate the trapped radiation. The system relies on three energy levels in the atoms (Fig. 1). The frequency of the trapped microwave radiation is close to, but not the same as the frequency of transitions between two of the states, *e* and *i*. The detuning is enough to ensure that the atoms cannot absorb any photons and make the transition from level *e* to *i*. Instead, the energy of state *e* is changed by a 'Stark' shift which increases linearly with the number of photons: each additional photon in the system of Brune *et al.* shifts the level by 400 kHz, far larger than the inherent widths of the levels or any other complicating factors.

The photon counting is done by passing these Rydberg atoms through the microwave cavity (Fig. 2). Just before entering it, each atom is excited, by radiation resonant with the *f*→*e* transition, into a superposition of these two states. When an atom emerges from the cavity, the component in level *f* remains unchanged, but the component in level *e* has had its phase changed by an amount proportional to the number of photons in the cavity and the time spent there (effectively, the atom's velocity). The atom is then bathed with more microwaves resonant with the *f*→*e* transition.

The result of the two bouts of excitation is a set of interference fringes in the probability of the atoms being found in level *e*. These fringes, called Ramsey fringes after their originator and the basis of the atomic clock, are due to interference between the amplitude for excitation in the two Ramsey zones.

Initially the field inside the cavity will typically not have a specific number of photons: instead it will be in a superposition of states with different numbers of photons. Brune *et al.* show that as the experiment progresses and more atoms pass through the cavity, interference modulates the field's photon probability distribution: in the authors' words, each interaction "decimates" the distribution. After a small number of such interac-

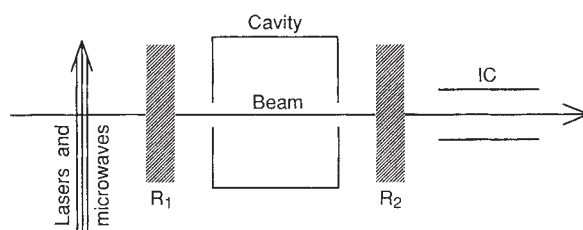


FIG. 2 The QND experiment of Brune *et al.*¹. Lasers and microwaves raise atoms in a beam into a high-lying Rydberg state. Microwaves in the first Ramsey zone (*R*₁) create a superposition of states, which then passes through the high-quality microwave cavity, whose field photons are to be counted. After further excitation in the second Ramsey zone (*R*₂), atoms are detected in an ionization counter (IC).

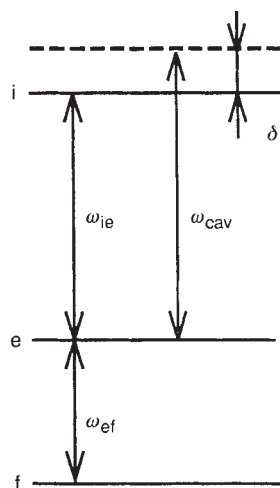


FIG. 1 Atomic energy levels in a quantum nondemolition experiment. The frequency of radiation in the cavity (ω_{cav}) is slightly different (by δ) from the resonant frequency (ω_{ie}) between two of the atomic states, but is close enough to shift that resonance by a small amount through the Stark effect; state *f* is too far away to be affected this way. Atoms in the QND experiment are put in a superposition of states *f* and *e* by microwaves tuned to ω_{ef} .

tions, the cavity field is forced into a pure number state picked at random from the initial distribution.

The fringes responsible for the QND 'decimation' equally reflect a change in the phase of the trapped field. If an atom, after interaction with the two Ramsey zones and the cavity field is detected in an excited state, we may directly infer the associated state of the field. An excited atom implies the existence of a superposition of two field states in the cavity, one of which is phase-shifted from the unchanged original. The next atom to be injected (probably with another velocity) and which survives to be detected in an excited state will cause each of these components in the field superposition to split further into two more phase-shifted states in a kind of phase diffusion. This phase diffusion is the signature of the formation of a pure number state.

But, and this is where Schrödinger's cat comes into the story, with a beam

1. Brune, M., Haroche, S., Raimond, J. M., Davidovich, L. & Zagury, N. *Phys. Rev. A* **45**, 5193–5214 (1992).
2. Braginsky, V. B. & Khalili, F. I. *Sov. Phys. JETP* **51**, 859–862 (1980).
3. Caves, C. M., Thorne, K. S., Drever, R. W. D., Sandberg, V. D. & Zimmerman, M. *Rev. Mod. Phys.* **52**, 341–392 (1980).
4. Levenson, M. D. *et al. Phys. Rev. Lett.* **57**, 2473–2476 (1986).
5. Meystre, P. in *Progress in Optics* (ed. Wolf, E.) (North Holland, Amsterdam, in the press).

containing atoms of only one velocity (perhaps obtained with laser cooling) the atoms can maintain the superposition of two phase-shifted fields; each interaction merely recreates the pre-existing superposition. Like Schrödinger's cat, the two fields are macroscopic states, and they are distinct states. Conventional Schrödinger cats rapidly die on interaction with a dissipative environment: their decay depends on how distinct are the components. But these cats are fed 'quantum food' in the form of the injected atoms in superposed states⁵.

Brune *et al.* show that their cats can be

made with existing experimental technology: the field damping time in cavities currently available is 0.1 seconds. Schrödinger cat states involving as many as 100 photons could be generated and observed, using a modified interference technique. If realized, they could form the basis of much-needed experimental investigation of the creation of quantum coherence and its destruction by the environment, a subject of intense theoretical speculation. □

Peter Knight is at the Blackett Laboratory, Imperial College, London SW7 2BZ, UK.

RNA CATALYSIS

The universe expands

Peter B. Moore

FOR over a decade, RNA enthusiasts have been claiming that a race of self-replicating entities that used RNA molecules both for genetic purposes and as catalysts for their metabolic reactions were the progenitors of us all. Their advocacy stems, of course, from the knowledge that some RNAs — ribozymes — possess catalytic activity. The chemistry catalysed by the ribozymes identified so far is not all that exciting, however, in that they promote nucleophilic attacks by either hydroxyl groups or water on phosphates. The faithless have questioned whether such a limited catalytic palette could sustain self-replicating entities, and so have cast doubt on the whole concept.

Two papers in last week's *Science*^{1,2} should silence some of the critics. They show that ribozymes can catalyse nucleophilic addition and elimination reactions involving carbonyl groups, and that a ribozymal activity of that sort is almost certainly responsible for the formation of peptide bonds during protein synthesis.

The peptidyl transferase activity of the ribosome is intrinsic to its large subunit³, and those in the business have known for several years that the Noller laboratory at Santa Cruz was trying to find out whether that activity is RNA-driven. The first formal report¹ on what has been accomplished is one of the two papers in question here.

The fragment reaction assay for peptidyl transferase activity⁴ is the key to this work. What is measured is the transfer to puromycin of amino acids esterified to an oligonucleotide isolated by nuclease digestion from the 3' end of an aminoacylated transfer RNA. Its virtue is that it can use isolated ribosomal subunits, independent of tRNA binding, messenger RNA binding, interactions with the small subunit, and all the other complications of protein synthesis.

Reconstitution studies carried out in the 1970s suggested that large ribosomal subunits from *Escherichia coli* containing less than their normal complement of protein retain fragment reaction activity⁵, and what Noller and his colleagues wanted to see was whether all the protein could be removed without loss of activity. Their best results were achieved using large subunits from the extreme thermophile, *Thermus aquaticus*. Proteins were stripped from them by treatment with sodium dodecyl sulfate (SDS), proteinase K and phenol (at 4° C). Done singly or in combination these treatments had no effect on activity, but any treatment that disrupted RNA tertiary structure abolished activity completely. The 50S subunits from *E. coli*, which presumably are not as tough, retained some activity in the face of SDS and proteinase K treatment, but would not tolerate phenol extraction in the cold. The activity remaining in stripped subunits had the same sensitivity to antibiotics as the activity that was seen in intact subunits.

These observations do not constitute absolute proof that the ribosomal peptidyl transferase is ribozymal because 5 per cent of the protein cannot be removed from *T. aquaticus* subunits by these means. It nevertheless suggests how bets should be placed in the future.

The second of the two papers, which also involves Noller, comes from Cech's laboratory at Boulder². It describes the conversion of the self-splicing intron from *Tetrahymena thermophila* from a catalyst of phosphodiester transesterifications into a catalyst of carboxylate ester hydrolysis. The guide sequence of the *Tetrahymena* nucleotide intervening sequence (IVS), which normally aligns its RNA substrates in the activity site by base pairing, was altered by mutation so that it would bind the 3' hexanucleotide

of tRNA^{fMet}. The site was engineered so that *N*-formyl-L-methionine esterified to the hexanucleotide's 3' ribose would occupy the active site. The mutated enzyme retains the ability to cleave substrate RNAs whose sequences have been altered so that they will bind to the mutated guide sequence, but now it also weakly catalyses the hydrolysis of the amino acid ester in question.

'Weakly' is very weak indeed — a five-fold increase in rate is all that is observed. Nevertheless, tests with variants of the IVS demonstrated that mutants lacking nuclease activity are also inactive as carboxylic acid esterases. In addition, the esterase activity depends on magnesium, as does the IVS nuclease activity, and it can be saturated by raising the substrate concentration. Thus this esterase activity is likely to be due to chemistry occurring at the normal active site of the IVS, an important point given the small effect observed.

The IVS esterase activity differs from the normal IVS transesterification activity in its nucleophile specificity. The latter accepts either guanosine (well) or water (poorly) as its nucleophile, but only water serves for the former. As the authors point out, the transesterification reactions of (tetrahedral) phosphate groups depend on in-line attack by nucleophiles. Carbonyl groups are trigonal, and nucleophiles attack them normal to the plane. Thus a guanosine positioned in the IVS active site for participation in a phosphate transesterification reaction will not be in a position to attack a carboxylate ester, if the carboxylate ester binds in the position normally occupied by phosphate groups.

Taken together, the two new papers deliver a simple message: we must expand our view of what RNAs can catalyse. That expansion is not a large one chemically; nucleophilic additions to phosphates are not all that different from nucleophilic additions to carbonyl groups. But it is of the utmost significance biologically. The type of RNA catalysis demonstrated in the papers is that on which the entire translation process depends. The possibility that there once were RNA entities capable of both replicating RNA and synthesizing protein seems a lot more reasonable now than it did a few months ago. □

Peter B. Moore is in the Department of Chemistry, Yale University, New Haven, Connecticut 06511, USA.

1. Noller, H. F., Hofferth, V. & Zimniak, L. *Science* **256**, 1416–1419 (1992).
2. Piccirilli, J. A., McConnell, T. S., Zaug, A. J., Noller, H. F. & Cech, T. R. *Science* **256**, 1420–1424 (1992).
3. Monro, R. E. *J. molec. Biol.* **26**, 147–151 (1967).
4. Monro, R. E. & Marcker, K. A. *J. molec. Biol.* **25**, 347–350 (1967).
5. Moore, V. G., Atchison, R. E., Thomas, G., Moran, M. & Noller, H. F. *Proc. natn. Acad. Sci. U.S.A.* **72**, 844–848 (1975).

Genetics and geography

N. Goldman and N. H. Barton

WHERE and how did the human species originate? Molecular genetic studies promise to give a definite answer, but at times they have flattered only to deceive. The controversy over human mitochondrial variation¹⁻⁶ shows how hard it will be to unravel the information contained in DNA sequences, and at present data are accumulating faster than our understanding of how to analyse them. Yet progress is being made — methods are now being developed for making rigorous distinctions between alternative hypotheses, and for estimating population sizes and rates of gene flow from sequence data⁷⁻⁹.

Our understanding of the genetic basis of evolutionary change has centred on following the frequency of different kinds of gene — so-called 'bean-bag genetics'. This approach has been challenged by the data now emerging on DNA sequence variation. Describing variation in terms of the frequency of A, T, G and C at individual base pairs would hardly be fruitful: because recombination between adjacent bases is very rare, even nuclear DNA is passed on in large blocks. The population is thus most naturally represented by the genealogical history of these blocks. Although most argument has been over statistical techniques for estimating genealogies, the relation between genealogies and populations is both more fundamental and harder to understand.

In 1987, Cann, Stoneking and the late Allan Wilson⁵ published data on variation in the mitochondrial (mt) DNA of people from different continents. Mitochondrial DNA was chosen for the study because it accumulates substitutions rapidly, and because it is passed on intact from mother to offspring, and so can be used to trace maternal ancestry without the complications of recombination. Cann *et al.* argued that the common ancestor of all these sequences — the so-called 'mitochondrial Eve' — was likely to have lived in Africa around 200,000 years ago. The surprise was not that all sequences traced back to one woman. Rather, the paper was important because it claimed to locate mitochondrial Eve in time and space, and to give information about the size and movements of the human population as a whole.

The results of Cann *et al.* were based on the application of parsimony analysis^{10,11}, a method originally designed for deducing the evolutionary relationships between species. Every possible tree is given a 'length', equal to the minimum number of evolutionary

changes required to explain the observed differences. The parsimony criterion says that the tree requiring fewest changes (the 'shortest tree') is preferred. Unfortunately, for problems involving large numbers of species (over 20 or so) there are too many trees for an exhaustive search and there is no guarantee that current computer programs will find the best tree; nor can they tell whether the shortest tree found is actually the best.

Cann *et al.* analysed the mtDNA of 147 individuals drawn from five geographical areas: Africans, Asians, Caucasians, native Australians and New Guineans. The presence or absence of 195 sites which could be cut by restriction enzymes defined 133 distinct haplotypes. The authors placed a geographical 'root' within the shortest tree they found, which required 312 mutational changes. The first split divided this tree into one group which contained only sequences from Africa, and another which consisted of a mixture of African and non-African sequences. Cann *et al.* then applied the parsimony criterion to argue that an African origin would require fewest migrations, and so was the best supported.

Doubts

There have been doubts about more than one aspect of these conclusions. Last year, D. Maddison¹ performed a thorough reanalysis and discovered that the tree was not in fact the best fit to the data: he found more than 10,000 trees which required only 307 mutations. Most of these did not indicate an African origin by the criterion used by Cann *et al.* Furthermore, Maddison pointed out that Cann and colleagues' single, non-optimal tree did not provide significant evidence for an African origin: an ancient migration into Africa was as parsimonious as one the other way. In reply, Wilson's group¹² could say little more than that parsimony could be misled by simple events at the population level, such as migration, when used to analyse intra-species variation.

Late last year, Wilson's group published another study. Vigilant *et al.*⁶ looked directly at variation in the control region of mtDNA sequence of 189 individuals from seven African populations and from Asians, Europeans, native Australians, New Guineans and African Americans. They again used parsimony to infer the evolutionary tree. Now, however, the root of the tree was located by including a chimpanzee in the analysis: in principle, using such an outgroup is more accurate than taking the root to

be at the midpoint, as was done before. Vigilant *et al.* found 100 equally parsimonious trees which again indicated an African origin; two statistical tests appeared to verify the significance of the findings.

Their findings have again come under fire, however, and for similar reasons. Templeton² and Hedges *et al.*³ both found many shorter trees, most of which did not suggest an African origin. In a longer paper, published this month, Maddison *et al.*⁴ reach the same conclusions; interestingly, they are the first group prompted to reanalyse the data by the "improbable" trees proposed. These three studies used different strategies to get the best out of the computer programs that search for the most parsimonious trees. No consensus has been reached in their papers or in the ensuing heated exchanges on electronic bulletin boards, but it is clear that the methods used by Wilson's group were inadequate, and so gave them a highly biased view of what the best trees were. Disturbingly, Maddison *et al.*⁴ also point out that the divergence between the chimp and human mtDNA is so much greater than that within our species that the chimp sequences used may be no more reliable for indicating the root of trees than a random sequence.

Parsimony is not universally accepted, even for the simpler case of inter-species studies. Setting aside the problems of finding the optimal tree, which is common to all such analyses, it is not clear that the shortest tree is the best estimate of the true tree¹³. Under some circumstances, parsimony gives the wrong tree with increasing certainty as more data are collected¹⁴. An alternative is to assume some explicit model of evolution, and use 'maximum likelihood': that is, choose the tree which would yield the observations with highest probability. Though computationally demanding, this approach has the advantages that it forces one to define an explicit statistical model, and allows straightforward comparisons of alternative hypotheses¹⁰. Moreover, as we argue below, it can be extended to include geographical information, and to give estimates of parameters describing the evolution of the whole population.

Most of the argument has centred on how to find the genealogy that best fits the data. But more fundamental is to know what the true tree would tell us about the location of the ancestors, and about movements of the whole population. Vigilant *et al.* proposed a "geographical states" method for testing the significance of an African origin: this is based on the probability that the 14 deepest branches in the tree would all lead to exclusively African groups, as they observed⁶.

RÉSUMÉ

Maddison *et al.*⁴ now point out several pitfalls in this procedure, which indicate general difficulties in building significance tests for genealogical data. First, although the test takes as a null hypothesis that the tree is constructed by the successive addition of African or non-African clades, rejection of this null model does not necessarily support one hypothesis about ancestral location over another. Second, the test takes the structure of the tree as certain, even though it is in fact estimated from the data; this could favour the preferred 'African Eve' hypothesis. Finally, the probability calculations were incorrect, in that they found the chance of seeing *precisely* the observed locations, rather than the chance of seeing that, plus all more extreme patterns. Given these difficulties, Maddison *et al.*⁴ instead used the straightforward criterion of parsimony to infer ancestral location, choosing that which requires fewest migrations. This procedure gave no clear evidence that our common mitochondrial ancestor lived in Africa; some trees suggested an African origin, some did not.

Worry

Human mtDNA sequences, like most such data, have been analysed using methods originally designed for establishing the relations between species. These methods tell us about the ancestry of one gene, but not about the structure of the whole population or about the forces that determine the spread of the gene we observe. A serious worry is that the relatively recent ancestry of human mitochondrial genomes reflects the recent spread of a favourable mutation, and so can tell us nothing about the population as a whole. Such 'selective sweeps' have reduced genetic diversity in regions of low recombination within the nuclear genome¹⁵.

If we take mtDNA to be strictly neutral, the probability of any particular genealogy being produced within a panmictic population of N genes can be written explicitly: at any time, the chance that two lineages will coalesce into one is just $1/N$. So one can use maximum likelihood to estimate the population size, and to test whether changes in the rate of coalescence of lineages reflect significant changes in population size, as in a founder event. Felsenstein has shown that this likelihood approach gives much more information than concentrating on the date of the one common ancestor, or on pairwise measures of sequence divergence⁷. When very many individuals are sampled, one-third of pairwise comparisons will be between lineages separated by the deepest split in the tree, and so will reflect one chance event. Maximum likelihood is more effi-

cient because it gives appropriate weight to all the branching events.

As the number of individuals increases, it becomes progressively harder to resolve the true tree³. But Felsenstein has shown how the likelihood approach extends naturally to give a sum over the set of all plausible trees⁸. Although this takes much more computing time than a search for the optimal tree, some such averaging over trees is essential. The problems with the interpretations of Wilson's group stemmed largely from concentration on one arbitrarily chosen tree — which is of course common practice.

Although Felsenstein only discusses the simple case of a single randomly mating population, it should be possible to extend likelihood methods to include hypotheses about the location of the ancestral population, and rates of gene flow. Slatkin and W. Maddison have already developed cladistic methods for estimating gene flow⁹. Their simulations reveal a consistent relation between the minimum number of migration events estimated by parsimony, and the actual number of migrants in each generation; they show that the tree of Cann *et al.*⁵ gives significant evidence of population subdivision.

Wilson's group pioneered the study of intra-specific variation in DNA sequence: ultimately, such variation must reveal more about the evolutionary processes that have shaped life than will comparisons between species. The controversy over the interpretation of geographical patterns in human mtDNA has shown the need for statistically sound methods of making inferences from genealogies, however, and the question of whether or not genetic data support an African origin either for our mitochondria, or for the human population as a whole, is still open. □

N. Goldman is in the Laboratory of Mathematical Biology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK. N. H. Barton is in the Division of Biology, University of Edinburgh, Edinburgh EH9 3JT, UK.

1. Maddison, D. R. *Syst. Zool.* **40**, 355–363 (1991).
2. Templeton, A. R. *Science* **255**, 737 (1992).
3. Hedges, S. B., Kumar, S., Tamura, K. & Stoneking, M. *Science* **255**, 737–739 (1992).
4. Maddison, D. R., Ruvolo, M. & Swofford, D. L. *Syst. Biol.* **41**, 111–124 (1992).
5. Cann, R. L., Stoneking, M. & Wilson, A. C. *Nature* **325**, 31–36 (1987).
6. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. *Science* **253**, 1503–1507 (1991).
7. Felsenstein, J. *Genet. Res.* **59**, 139–147 (1992).
8. Felsenstein, J. *Genet. Res.* (in the press).
9. Slatkin, M. & Maddison, W. P. *Genetics* **126**, 249–260 (1990).
10. Felsenstein, J. A. *Rev. Genet.* **22**, 521–565 (1988).
11. Swofford, D. L. & Olsen, G. J. in *Molecular Systematics* (eds Hillis, D. M. & Moritz, C.) 411–502 (Sinauer, Sunderland, Massachusetts, 1990).
12. Wilson, A. C., Stoneking, M. & Cann, R. L. *Syst. Zool.* **40**, 363–365 (1991).
13. Goldman, N. *Syst. Zool.* **39**, 345–361 (1990).
14. Felsenstein, J. *Syst. Zool.* **27**, 401–410 (1978).
15. Charlesworth, B. *Nature* **356**, 475–476 (1992).

Light work

THE art of manipulating atoms in beams of light emerges from the realm of the esoteric into the plainly practical with the description of a scheme for directly drawing electronic structures on silicon substrates (M. Prentiss *et al.* *Appl. Phys. Lett.* **60**, 1027–1029; 1992). Conventional photolithography works by using light to transfer a pattern (that will ultimately become the microcircuit) from a material mask to the silicon substrate. Prentiss *et al.* reverse matters: an atom beam directed at the substrate will have to pass through focused laser beams to be directed into place on the surface. In their experiment, lines a few tenths of a millimetre across are written onto the silicon surface using a beam of sodium atoms and orange laser light. The ultimate prospect, the authors hope, is for direct writing with the resolution of a single wavelength of light.

Dog's dinner

It is one thing to show that oral vaccination of foxes in the wild may stem the tide of rabies in Europe (*Nature* **354**, 520–522; 1992), quite another to find a bait attractive to potential carriers of the disease elsewhere. H. Kharmachi *et al.* now report on their trial, with biomarkers, of the comparative allure of sausage, fishmeal, chicken head and sponge baits to Tunisian dogs (*Vet. Rec.* **130**, 494; 1992). The sponges contained "an attractant of minced meat, eggs, yoghurt, fish and cheese which had been kept for several days at room temperature", but came a poor last in the acceptance stakes at 30%. Chicken heads (98%) were most attractive. Chicken head and vaccine sachet would have to be brought together by hand, however, making the bait time-consuming to prepare.

Bright spark

STARS crushed by black holes may be the source of γ -ray bursters, suggests B. Carter (*Astrophys. J.* **391**, L67–L70; 1992). If, as seems likely in the light of data from the Compton observatory, these flashes of γ -radiation originate outside our Galaxy, they must be exceptionally energetic. For several years Carter has been studying the fate of stars that come too close to black holes, but has assumed that little of the inevitable heating comes out as radiation. What, he asks now, if that energy is converted into γ -rays? The doomed stars are compressed into pancakes and raised to temperatures of 10^9 K. The radiating phase could last a fraction of a second to many tens of seconds, in the range actually seen. Although he says the proposal is speculative, Carter highlights the possibility of identifying what kind of star has been consumed from the way the γ -ray intensity rises and falls.

Beads, bacteria and actin

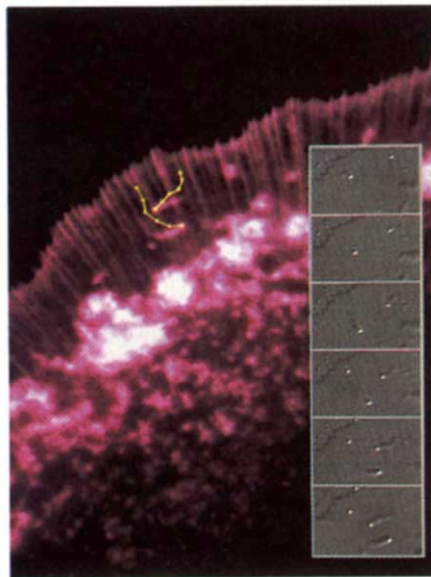
Joseph W. Sanger and Jean M. Sanger

ONCE again beads have been used to trick Mother Nature into revealing how actin moves things around on the cell surface^{1,2}. On page 515 of this issue², Forscher and his collaborators illustrate how polycationic beads, which bind to the dorsal surface of nerve cell growth cones, can induce actin filament bundles, termed inductopodia, to form inside the cells. Not only can the extracellular beads induce intracellular actin filament assembly, but the beads associated with the filament bundles can then move along the cell surface. This striking phenomenon can be appreciated best by video-enhanced microscopy, as the authors showed at the meeting of the American Society for Cell Biology in December of last year.

Actually, two types of movement can be observed when beads are applied to the growth cone cell surface: a retrograde motility (away from the growth cone towards the nucleus), at the rate $1\text{--}6\ \mu\text{m min}^{-1}$, that is coupled to the submembranous, filamentous (F-) actin network¹; and the newly identified motility coupled to the assembly of inductopodia (see figure)². The trajectories of beads associated with elongating inductopodia can be straight or curved, and can be at any angle to retrograde flow.

The highest velocities occur when the beads move parallel to and in the same direction as retrograde flow. Subtraction of the retrograde F-actin flow rate from the total rate of bead motility shows that the bead-induced rate averages $10\ \mu\text{m min}^{-1}$ regardless of the angle of the path of movement. A requirement for continuing actin polymerization was shown by treating the beaded cells with cytochalasin, which disrupts F-actin. Within 15 seconds inductopodia elongation ceased, and both beads and inductopodia moved rearward in synchrony with the retracting cortical F-actin. As soon as cytochalasin was washed out, inductopodia growth and bead motility resumed. Coating the polycationic beads with a non-specific protein such as albumin also inhibited inductopodia formation but not the bead motility associated with retrograde flow.

The similarity of the bead-induced motility to that of some infectious bacteria such as *Listeria*^{3,4} and *Shigella*^{5,6} is as striking as the bead motility itself. These bacteria are engulfed by host cells but escape an unpleasant fate in the acidic environment of the phagolysosome by entering the infected cells' cytoplasm. In both systems, bead and bacterium, actin is recruited to polymerize into



Polycationic beads in motion on the surface of a growth cone cell. The electronically enhanced video sequence (right, 10-second intervals) illustrates the movements of 240-nm-diameter beads on the membrane surface. The two lower beads trigger submembranous actin filament assembly which is tightly coupled to lateral bead translocation. The background image shows the F-actin structure in this growth cone, demonstrating that the comet tail-like protrusions behind beads result from accumulation of F-actin. Corresponding bead positions over time are marked in yellow for the two lower beads in the video sequence. (Illustration courtesy of P. Forscher.)

tail-like appendages^{6,7} and motility is inhibited reversibly by cytochalasin³. The trajectories and rates of motility are also remarkably similar^{3,4}. The actin filaments in the bacterial tails are arranged with their fast-growing, or barbed ends, proximal to the bacteria⁸, and it is at this juncture where actin monomers insert as the tails elongate⁴. When a region of the tail is marked with fluorescent actin, the marked region remains stationary as the bacterium moves forward⁹. These data are consistent with a model in which the force of actin polymerization propels the bacteria forward^{3,4}, although a role for myosin has not been ruled out entirely. The questions of both actin filament polarity and the actual site of actin polymerization are yet to be determined for inductopodia.

Whether or not actin-nucleating proteins are involved in either system is not known; such proteins promote the polymerization of actin monomers into filaments. But whereas the bacteria might synthesize them, clearly the beads do not. The binding of the beads to the

cell membrane must initiate a transmembrane signal that enables nucleating proteins to localize at the binding site. Is it possible that, once they have entered the host cell, the bacteria in some way mimic the transmembrane complex and compete for the cells' actin-nucleating proteins?

The new paper by Forscher *et al.*² may also point to a fresh way to look at another group of infectious bacteria that can histopathologically insult our intestines. When strains of enteropathogenic *Escherichia coli* attach to the intestine, the microvillous lining, so important in absorption, is lost in a process termed effacing¹⁰. It is not known how this process occurs. The bacteria normally do not enter the cell but are thought to remain firmly attached to the cell membrane. There is intriguing evidence that tufts of filaments project into the cytoplasm from the attachment sites¹⁰. Could these filaments be actin? And is it possible that these dysentery-causing bacteria move along the intestinal cell surfaces inducing the loss of the microvilli by appropriating their supply of actin? Cell biologists with intestinal fortitude might look into these questions.

Understanding how actin and other proteins are recruited to a particular region of the cell's plasma membrane remains a tough problem. This process is central to the formation of cleavage furrows, synaptosomes, myofibrils and stress fibres, and the bead experiments by Forscher *et al.*² suggest a way to analyse it — isolation of the beads with their inductopodia or *Listeria* with their actin tails may provide the means to isolate and identify the proteins involved. It is probably healthier to work with beads, but anyone who looks at *Listeria* travelling through infected cells will soon be seduced into considering the bacterial system. □

Joseph W. Sanger and Jean M. Sanger of the Pennsylvania Muscle Institute are in the Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19194-6058, USA.

1. Forscher, P. & Smith, S. J. *J. Cell Biol.* **107**, 1505–1516 (1988).
2. Forscher, P., Lin, C.H. & Thompson, C. *Nature* **357**, 515–518 (1992).
3. Dabiri, G.A., Sanger, J.M., Portnoy, D.A. & Southwick, F.S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6068–6072 (1990).
4. Sanger, J.M., Sanger, J.W. & Southwick, F.S. *Infect. Immun.* (submitted).
5. Ogawa, H., Nakamura, A. & Nakaya, R. *Jap. J. med. Sci. Biol.* **21**, 259–273 (1968).
6. Bernardini, M.L., Mounier, J., d'Hauteville, H., Coquis-Rodon, M. & Sansonetti, P. *J. Proc. natn. Acad. Sci. U.S.A.* **86**, 3867–3871 (1989).
7. Tilney, L.G. & Portnoy, D.A. *J. Cell Biol.* **109**, 1597–1608 (1989).
8. Tilney, L.G. *et al. J. Cell Biol.* (in the press).
9. Theriot, J.A., Mitchison, T.J., Tilney, L.G. & Portnoy, D.A. *Nature* **357**, 257–259 (1992).
10. Helle, P., Morin, M., Jacques, M. & Fairbrother, J.M. *Infect. Immun.* **59**, 814–821 (1991).

C₆₀ chemistry expands

Leonard F. Lindoy

NEW milestones in the emerging chemistry of C₆₀ are described in this issue by H. W. Kroto and coworkers at the University of Sussex (*Nature* **357**, 479–481; 1992), and in reports by teams at Sandia National Laboratories and Toyohashi and Mie Universities.

Until recently, the main focuses of C₆₀ research have been the structural, physical and theoretical aspects of this remarkable molecule. However, following the development of procedures for obtaining usable quantities of pure C₆₀ increased emphasis is being given to its chemical properties. Chemical studies on the less abundant fullerenes, such as C₇₀, are also starting to appear.

The excitement and challenge felt in exploring the emerging chemical reactivity of the fullerenes resembles that which occurred during the development of aromatic chemistry last century, following Kekulé's description of the structure of benzene in 1865. Chemically modified fullerenes open the way to a virtually endless variety of new materials. The potential for the future application of these modified fullerenes is clearly very great when one considers the extraordinary properties (including superconductivity, optical properties and ferromagnetism) of the parent icosahedral cage and its doped derivatives. The likelihood that the properties of a given derivative will be able to be matched precisely to the needs of a particular application via 'chemical tuning' also has far reaching implications.

C₆₀ is composed of tri-coordinated carbon atoms making up both 5-membered and 6-membered rings on the surface of the cage. The inherent stability of the cage reflects its benzenoid nature. Although all the carbon atoms are equivalent, there are formally two C–C bond types present in C₆₀: those that are common to two fused 6-membered rings and those common to a 5-membered ring fused with a 6-membered ring. Further, the π -electron density is not evenly delocalized throughout the cage, because resonance structures incorporating double bonds in the 5-membered rings are unfavourable.

The non-equivalence of bonds throughout C₆₀ is expected to be reflected in reactivity patterns which differ from classical aromatic compounds. Indeed, clear reactivity differences are emerging as the chemistry of the fullerenes unfolds. Nevertheless, much further work is required to establish fully the reactivity ground rules that apply to the new cages (and their derivatives).

Various derivatives containing atoms

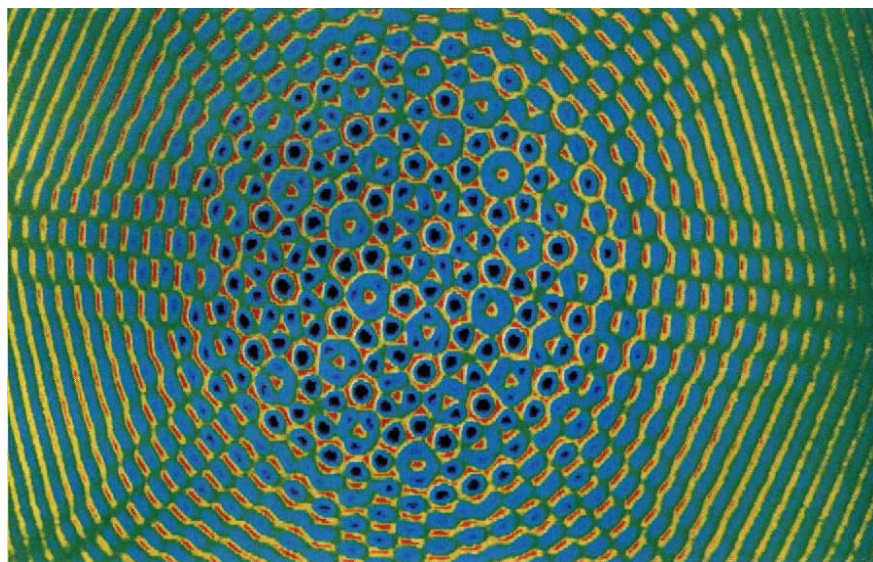
or groups covalently bound to the outside of the C₆₀ cage have been synthesized but in most cases neither the relative positions of the appended groups nor the isomeric purity of the respective products has been determined. Typical solid products in this category produced in an 'uncontrolled' manner include C₆₀H₃₆, C₆₀F₃₆, and C₇₀F₃₄. Related chlorinated and brominated derivatives have also been obtained. Although the nature of these products has not been fully defined, the integrity of the cage framework appears to be maintained in each compound.

In the work of Kroto and colleagues, the new brominated species C₆₀Br₆ and C₆₀Br₈ have been prepared and the sites of bromination precisely determined using X-ray diffraction (see pages 480 and 481). These products were obtained under closely defined conditions by the reaction of solutions of C₆₀ with bromine. A third product, C₆₀Br₂₄ was also isolated from the reaction of C₆₀ with bromine in the absence of solvent. This

latter product was also reported recently by F. N. Tebbe *et al.* (*Science* **256**, 822–825; 1992) who determined its X-ray structure. The 24 bromine atoms envelop the carbon core in a symmetrical fashion, shielding the 18 remaining double bonds from further addition. Because halogenated C₆₀ species have been demonstrated to act as synthetic intermediates, the availability of the new, well characterized and isomerically pure derivatives represents a significant advance — especially in promoting the development of the systematic chemistry of the fullerenes.

The tendency for six or eight groups (or a larger number of groups which are low multiples of these values) to add to C₆₀ is now documented for a number of different reaction types. Such a realization provides a starting point for unravelling the interplay of electronic influences that clearly underwrite the chemistry of this cage.

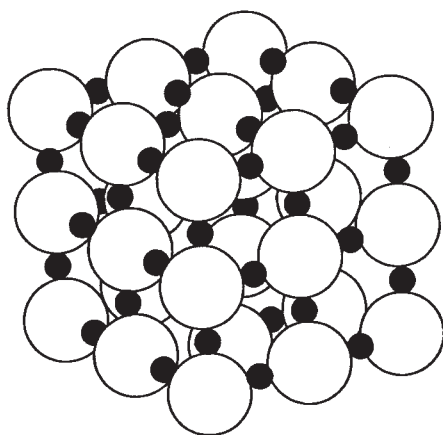
Following reports that C₆₀ readily reacts with benzylic radicals to form polybenzylated fullerenes, A. D. Loy and R. A. Assink at Sandia used the related diradical from xylylene in the synthesis of a C₆₀-*p*-xylylene copolymer in which individual C₆₀ molecules are



THE surface of liquid ethanol is an unlikely place to look for insights into two-dimensional crystals. But B. Christiansen *et al.* (*Phys. Rev. Lett.* **68**, 2157–2160; 1992) have shown that the ripples generated by simply shaking a shallow trough of the liquid can adopt a wide range of ordered patterns.

The patterns are the result of interactions between standing capillary waves (surface waves controlled by surface tension). Christiansen *et al.* generate relatively long-wavelength capillary waves by vertical vibration of a cylindrical container, 8.4 cm in diameter. Transi-

tions from one pattern to another can be induced by changing either the frequency or the amplitude of vibration. When three standing waves interact, the pattern has hexagonal (honeycomb) symmetry; increasing the amplitude creates a square pattern from two standing waves. At lower frequencies, this pattern develops dislocations and ultimately breaks down into a triangular pattern without long-range order. The quasicrystalline pattern shown above is the result of interactions between four standing waves, and appears spontaneously from a largely disordered state. P.B.



The structure proposed by Nagashima *et al.* for the three-dimensional polymer, $(C_{60}Pd_3)_n$. The large open spheres each represent a C_{60} group, the palladium atoms are represented by black spheres.

covalently linked (*J. Am. chem. Soc.* **114**, 3978–3980; 1992). The product, which is believed to be extensively cross-linked, represents the first example of a new class of polymeric materials. Unusual properties are expected but no details are yet available.

The first organometallic polymers incorporating C_{60} units bridged by palladium atoms were described recently by Nagashima *et al.* (*J. chem. Soc., Chem. Commun.*, 377–379; 1992). The polymers are of type $(C_{60}Pd)_n$, $(C_{60}Pd_2)_n$ and $(C_{60}Pd_3)_n$. The authors postulate that the products are made up of alternating uncharged palladium atoms and carbon cages in regular arrays, with the palladium atoms serving to link the individual C_{60} groups in respectively one, two and three dimensions. The palladium appears to act as a bridge by binding to the π -electron density on the surfaces of two C_{60} molecules; precedence for such bonding already exists in the growing organometallic literature of C_{60} derivatives. The two-dimensional polymer has not been isolated but is considered to be an intermediate in the formation of the three-dimensional product. The latter (see figure) is clearly the most thermodynamically stable and, for example, may be generated by thermal disproportionation of $(C_{60}Pd)_n$.

In the presence of excess palladium, a product was obtained of stoichiometry $(C_{60}Pd_{3.5})_n$ which is believed to contain the excess palladium deposited on the surface of the polymer. This metal-rich product was observed to catalyse the heterogeneous hydrogenation of diphenylacetylene in cyclohexane — a harbinger for the importance of this class of polymer in the future. □

Leonard F. Lindoy is in the Department of Chemistry and Biochemistry, James Cook University, Townsville 4811, Queensland, Australia.

Sex and the male stick insect

Piero P. Giorgi

STICK insects are better known for their mimetic ability than for their contribution to our understanding of reproductive biology. During the day they remain immobile on branches, where they generally escape detection because their bodies bear an uncanny resemblance to sticks or leaves (Fig. 1). During the night they slowly move about feeding, and reproducing — often, as a group at the University of Bologna has found^{1–4}, in decidedly peculiar ways.

There are many variations on the theme of sexual reproduction. An egg can develop without activation by a sperm (parthenogenesis), or it can eliminate the male pronucleus after fertilization (gynogenesis). The male genetic material, however, is much less versatile and development in the absence of the female pronucleus (androgenesis) has so

(Fig. 2c); thus the hybridogenetic form is a perpetual F_1 hybrid, which can evolve to a full self-sustaining species if it shifts to parthenogenesis. Males of the two hybridogenic strains discovered by Scali and co-workers (*B. rossius-grandii benazzii* and *B. rossius-grandii grandii*) are infertile, whereas females, besides maintaining the hybridogenetic strain, can also reproduce by gynogenesis. Hybridogenesis has often been advocated as a possible mechanism of speciation in fish and amphibians⁶, but no case had yet been found in invertebrates.

The female stick insects had more surprises in store. When a female of *B. rossius-grandii benazzii* is fertilized by a *B. rossius* male, up to 20 per cent of the offspring have the genetic make-up of the father alone (Fig. 2e). This was first demonstrated by use of enzyme elec-



FIG. 1 Female of the stick insect *Bacillus rossius-grandii benazzii*.

far been described only in plants. Since 1887, when the Hertwigs first demonstrated experimental androgenesis in echinoderms⁷, students of animal biology have been taught that this phenomenon occurs only when induced artificially. But now the group at Bologna has discovered natural androgenesis, as well as other unexpected forms of reproduction, in certain species of stick insect.

Using gel electrophoretic separation of alloenzymes and chromosomal analysis to reconstruct the evolutionary history of the stick insect population of Sicily, Scali, Mantovani and Tinti identified two hybridogenetic strains which have arisen naturally from hybridization between *Bacillus rossius* females and *B. grandii* males of different subspecies^{1,2}.

Hybridogenesis is a type of reproduction particular to interspecific hybrids, and involves the maintenance of one parental genome in the gametes and the elimination of the other. The discarded genome is provided anew through fertilization by the host fathering species

trophoresis, and has now been confirmed by chromosomal analysis⁴. The genetic material of the mother is eliminated so that her eggs develop entirely under the instruction of paternal chromosomes. The resulting androgenetic males and females are fertile, and diploid through duplication of the active chromosomes, or by fusion of two sperm

1. Mantovani, B. & Scali, V. *Invert. Reprod. Develop.* **18**, 185–188 (1990).
2. Mantovani, B., Scali, V. & Tinti, F. *J. evol. Biol.* **4**, 279–290 (1991).
3. Scali, V., Mantovani, B. & Tinti, F. *Frustula Entomol. NS* **12**, 103–108 (1989).
4. Mantovani, B. & Scali, V. *Evolution* **46**, 783–796 (1992).
5. Hertwig, O. & Hertwig R. *Jen. Zeitschr. Med. Natur.* **20**, 120 (1887).
6. Uzzell, T., Holtz, H. & Berger, L. *J. exp. Zool.* **214**, 251–259 (1980).
7. Dubois, A. & Gunther, R. *Zool. Jahrb. Syst.* **109**, 290–293 (1982).
8. Mayr, E. *Animal Species and Evolution* (Harvard University Press, Cambridge, Massachusetts, 1963).
9. Paterson, H.E.H. in *Species and Speciation* (ed. Vrba, E.S.) 21–29 (Transvaal Museum Monograph No. 4, Pretoria, 1985).
10. Paterson, H.E.H. *Pacific Sci.* **42**, 65–71 (1988).
11. Darlington, C. D. *Evolution of Genetic Systems* (Oliver & Boyd, Edinburgh, 1958).

pronuclei. Females of the hybrid *B. rossius-grandii benazzii* cannot reproduce by parthenogenesis, suggesting that perhaps androgenesis could be a strategy to cope with complex requirements for nucleocytoplasmic interactions in hybrid development.

The discovery that androgenesis can occur naturally in animals opens new prospects in evolutionary and developmental biology. It has been pointed out that hybridogenesis is, practically, a case of sexual parasitism, where females kidnap and exploit sperm from other species. In fact the term 'klepton' (from the Greek for 'thief') has been proposed to denote a systematic group of hybridogenetic origin⁷. But Scali *et al.* identify cases of androgenesis when *B. rossius-grandii benazzii* hybrid females mated with males of *B. rossius* and, in some instances,

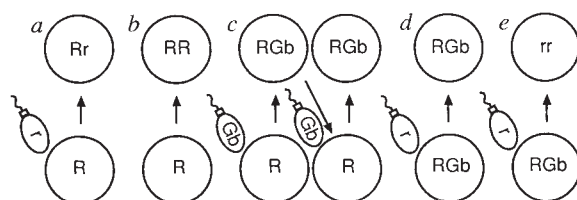


FIG. 2 Five types of reproduction found in stick insects. Bottom row, egg and sperm cells; top row, zygotes. R, female genome of *B. rossius*; r, male genome of *B. rossius*; Gb, genome of the subspecies *B. grandii benazzii*; RGb, genome of the hybrid *B. rossius-grandii benazzii*. a, Normal sexual reproduction; b, parthenogenesis; c, hemiclinal hybridogenesis; d, gynogenesis; e, androgenesis. In hemiclinal hybridogenesis, females of *B. rossius-grandii benazzii* (Fig. 1) produce eggs with only the R genome (oblique arrow).

even with *B. grandii benazzii*. The production of offspring corresponding to the original male species through androgenesis may now rehabilitate the hybridogenetic female from the status of a thief's daughter to that of a potential genetic store for the fathering taxon.

Finally, the existence of these complex modes of reproduction casts doubt on Mayr's 'isolation concept' of speciation⁸

while supporting the more recent 'recognition concept' proposed by Paterson⁹. This second concept views speciation in terms of specific mate recognition rather than reproductive isolation. Both hypotheses are restricted to sexual organisms. But in the recognition concept, parthenogenesis is considered as an escape from hybridity¹⁰, as originally suggested by C.D. Darlington¹¹, whereas Mayr and others see problems of hybridity in terms of isolation mechanisms.

Androgenesis, which also amounts to an escape from hybridity, would be expected to occur under the recognition concept, but is contrary to the expectations of the isolation concept. So the reproduction of stick insects can be explained better in terms of specific behavioural or molecular signals mediating recognition between individuals or gametes, rather than in terms of teleological barriers to mating. □

Piero P. Giorgi is in the School of Anatomy, University of Queensland, Brisbane, 4072 Australia.

CLIMATE CHANGE

Packing away carbon isotopes

Robin Keir

PACKRATS in the American southwest may have been unwitting assistants over the millennia in the modern attempt to understand climate change. Ancient desert shrubbery buried by the rats is analysed on page 461 of this issue¹, by Marino *et al.*, for the ratio of the carbon isotopes ¹³C and ¹²C. The interest of Marino *et al.*, and also of Leuenberger *et al.*² who report analyses of CO₂ trapped in ancient Antarctic ice (page 488), is in determining the ratios during the last ice age in the sea surface and below.

Since it was learned that the atmospheric CO₂ content during the ice ages can be 30 per cent lower than in warm climates, a great deal of effort has gone into reconstructing the glacial distribution of the ¹³C/¹²C ratio within the ocean. This is because the biogeochemical processes that transfer CO₂ from the surface to deeper waters also fractionate the carbon isotopes, so it is widely suspected that alterations in the dynamics of the ocean's geochemistry have been responsible for changes in atmospheric CO₂ and so tie into climate change.

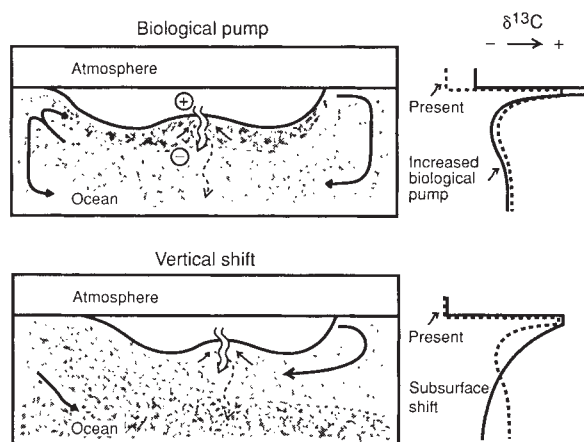
Until now, the ^δ¹³C isotope ratio of atmospheric CO₂, which is needed to pin down the sea-surface ratio in the last ice age has been uncertain, the few older ice-core measurements available being compromised by long storage times and other factors. The new results indicate that glacial atmospheric ^δ¹³C was prob-

ably 0.3–0.7‰ lower than in the modern, pre-industrial period. Given that artefacts could bias either approach, the two results agree well.

The numbers seem small but they are important. Air-sea exchange fractionates CO₂, leaving the atmospheric ratio 8.5‰ lower than the sea-surface ratio. The fractionation is dependent on sea-surface temperature, becoming stronger in colder waters. Leuenberger *et al.* point out that with the sea surface 2 °C cooler on average (as in the last ice age) the air-sea ^δ¹³C difference would increase by about 0.2‰. Thus the 0.3–0.7‰ decrease found for the atmospheric ^δ¹³C implies that the glacial surface ocean was 0.1–0.5‰ lower than in recent preindustrial times. Analysis of benthic foraminifera from deep-sea sediments indicate that the glacial deep ocean was on average 0.5‰ lower than at present³, so the atmospheric results appear to confine the increase in ^δ¹³C in the surface deep-ocean contrast during the last ice age to no more than 0.4‰, maybe less.

This makes it difficult to account for the decrease in

atmospheric CO₂ by either of two general types of carbon fractionating mechanisms that have been proposed, the biological pump and vertical shifts of dissolved carbon, for example through changed ocean circulation (see box and figure). Model calculations⁴ of increases in biological-pump efficiency indicate that each additional ‰ difference between the sea-surface and deep-sea ^δ¹³C corresponds to a decrease of about 100 parts per million (p.p.m.) in the concentration of atmospheric CO₂ (which stood at 280 p.p.m. before the industrial revolution). Various vertical mechanisms result^{5,6} in atmospheric CO₂ decreasing by 40–150 p.p.m. for each additional ‰ difference in ^δ¹³C. But the atmospheric CO₂ concentration was about 90 p.p.m. lower in glacial times than in the warm



Concentration of nutrients and ¹²C below the ocean surface (indicated by stippling) occurs according to the strength of the biological pump and pattern of ocean circulation.

Carbon isotopes and carbon cycling

BIOGEOCHEMICAL processes that can change atmospheric CO_2 and the ocean-atmosphere distribution tend to fit into two categories (see figure): biological-pump increases which involve greater sinking fluxes of organic carbon relative to upwelling of total CO_2 ; and subsurface rearrangement, where a change in ocean circulation or vertical distribution of organic matter re-oxidation actuates a transfer of nutrients and total CO_2 to greater depths⁵. Processes producing this vertical shift are sometimes called alkalinity mechanisms, because the downward shift of metabolic CO_2 , currently concentrated at intermediate depth, produces a transient increase in calcium carbonate through dissolution at the ocean floor, and it is this characteristic which reduces the atmospheric CO_2 . But ocean alkalinity also increases in response to an increased biological pump as well as to decreases in biological production of calcium carbonate at the surface. In the case of decreased carbonate production, the ocean isotope composition is virtually unchanged, and one cannot in-

fer changes in ocean alkalinity from the $\delta^{13}\text{C}$ distribution.

Both biological-pump and subsurface-rearrangement processes preferentially transfer ^{12}C over ^{13}C to the deep ocean and increase the $\delta^{13}\text{C}$ gradient from surface to deep ocean. The main difference is in where the ^{12}C is shifted from. Increases in the biological pump further deplete the light isotope from the surface ocean and the atmosphere because of gas exchange. Simple subsurface-rearrangement mechanisms, such as deeper recycling of particulate organic matter, transfer the ^{12}C exclusively from intermediate to deep waters, and the isotope ratio in the surface ocean and atmosphere is unaffected. In models with a polar outcrop of cold water, a subsurface rearrangement due to increased upwelling in the upper ocean may transfer light carbon out of the surface and intermediate waters, because the warm-ocean biological pump of ^{12}C from the surface becomes stronger in relation to the 'leak' of light carbon out of the cold surface ocean and back into the warm⁶. **R.K.**

preindustrial climate, more than twice what can be accounted for by the 0.4‰ change in surface-deep-ocean carbon isotope gradient.

One way out of this fix is if the ice-age air-sea exchange fractionation is more effective than expected from average sea-surface temperature. Then the atmospheric results allow a more positive ocean surface $\delta^{13}\text{C}$ and a greater vertical ocean gradient. Marino *et al.* show that this works if more of the gas exchange occurred through the cold polar ocean, especially if these waters decreased their $\delta^{13}\text{C}$ by as much as 1‰, as indicated by measurements⁷ from the planktonic foraminifera, *Neoglobobulimina pachyderma*. This decrease would be consistent with decreased productivity in southern-ocean waters inferred from opal accumulation rates and Ge/Si ratios⁸.

In the model of Marino *et al.*, an increase in the low-latitude biological pump, in the exchange of upper ocean water and a near cessation of high-latitude productivity apparently combine to produce a decrease in atmospheric CO_2 and a change in ocean-atmosphere $\delta^{13}\text{C}$ similar to that deduced from glacial proxies. But as loss of productivity would allow high CO_2 partial pressure to surface in the surface ocean, it is not clear how exhalation of isotopically light CO_2 into the atmosphere would decrease atmospheric $\delta^{13}\text{C}$ without pumping the atmospheric CO_2 content back up again.

The differing $\delta^{13}\text{C}$ values at the sea surface and in deep waters are a result of

the differing roles of dissolved CO_2 and of organic matter produced by phytoplankton. The reduced partial pressure of CO_2 during the last ice age, it has been suggested⁹, would alter the fractionation by biotic processes, invalidating the constant organic contribution built into the model calculations. Leuenberger *et al.* show that this variation could offset some of the increased $\delta^{13}\text{C}$ gradient expected from the simple biological pump and subsurface rearrangement mechanisms.

Deciphering the mechanism of atmospheric CO_2 change from the changes in carbon isotope distributions is a difficult game using small numbers, painstakingly obtained. What is evident from the results in this issue is that it is hard to explain the low ice age CO_2 solely using carbon-fractionating mechanisms. □

Robin Keir is at GEOMAR, Christian-Albrechts University, 2300 Kiel 14, Germany.

1. Marino, B. D., McElroy, M. B., Salawitch, R. J. & Spaulding, W. G. *Nature* **357**, 461–466 (1992).
2. Leuenberger, M., Siegenthaler, U. & Langway, C. C. *Nature* **357**, 488–490 (1992).
3. Curry, W. B., Duplessy, J. C., Labeyrie, L. D. & Shackleton, N. J. *Paleoceanography* **3**, 317–341 (1988).
4. Broecker, W. S. & Peng, T.-H. *Radiocarbon* **28**, 309–327 (1986).
5. Boyle, E. A. *J. geophys. Res.* **93**, 15,701–15,714 (1988).
6. Keir, R. S. *Global biogeochem. Cycles* **5**, 351–358 (1991).
7. Labeyrie, L. D. & Duplessy, J. C. *Palaeogeog. Palaeoclim. Palaeoecol.* **50**, 217–240 (1985).
8. Mortlock, R. A. *et al.* *Nature* **351**, 220–223 (1991).
9. Rau, G. H., Takahashi, T. & Des Marais, D. J. *Nature* **341**, 516–518 (1989).

DAEDALUS

Water music

MATTER absorbs light on strict quantum principles. A molecule absorbs a photon most readily if it has just the right energy to raise that molecule to some allowed excited state. So, says Daedalus, a liquid molecule should be neatly evaporated by a photon which exactly supplies its latent heat of evaporation. For water at room temperature, for example, this 'resonant evaporation' should be induced by photons of wavelength 2.71 micrometres, in the infrared. Water indeed absorbs strongly in this region.

Resonant evaporation, says Daedalus, has absolutely no thermal inertia. The photon must evaporate the receiving molecule instantly. A modulated infrared laser, tuned to 2.71 micrometres and aimed at a water surface, would vaporize it in precise synchronism to that modulation. Water expands strongly on evaporation; so pressure pulses exactly following the modulation would propagate into the atmosphere. If the modulation was in the audio band, the water would speak up in high fidelity.

The result is a loudspeaker with no moving parts. Its frequency extends from the deepest bass far into the ultrasonic region. Best of all, it generates its sound from an exact location — the point where the beam enters the water surface. At a stroke, it solves the whole problem of stereo reproduction.

A conventional two-track stereo recording is only an approximation of the original multi-track master recording, on which each instrument has its own track. But 'resonant evaporation stereo' could reproduce the master exactly. A bank of infrared lasers facing a damp wall, each driven from one instrumental track and aimed at the true position of that instrument, could bring the original concert to life in full spatial detail on that wall. (A better wetted surface would be one of those polymeric fabrics which can absorb vast amounts of free water without getting sticky.)

The hi-fi industry, in its ceaseless quest for expensive new gadgets, should welcome this complicated advance. So should the cinema. Each actor, tracked automatically and invisibly around the new damp screen by his own sonic infrared laser, would speak consistently from his own larynx. The stage could benefit too. Ventriloquists with damp dummies and pop stars with damp shirt fronts could engage in convincing double-talk or mime to their own records, with the aid of a confederate aiming an infrared laser from the balcony. Daedalus is even inventing laser-based aqua-karaoke. He reckons that a singer in the bath might rather enjoy the experience of having the bath join in. **David Jones**

Recent inactivity in African rift

SIR — Within the past decade, global positioning geodesy (GPS) has been aimed at a host of tectonic targets ranging from volcanic and seismic hazards to plate motions and global sea-level studies¹. These experiments typically reward their investigators when signals emerge above the 1-cm uncertainty commonly associated with GPS data. Thus plate-motion studies gain credibility after a few years of measurement, global sea-level studies after perhaps a few decades, while volcanic eruptions and earthquakes yield reportable results as and when they occur within existing geodetic networks. In short, geophysicists have set their GPS nets and now wait for the Earth to move.

Tectonic passivity is rarely of note. Yet in zones where dramatic volcano-tectonic architecture was formed yesterday, geologically speaking, the exception becomes interesting. The current widening rate of the East African rift has been estimated by various authors at 3–8 mm per year based on differing geological reconstructions². Sinistral shear amounting to 50% of the opening rate has also been proposed (ref. 3 and W. Hay, personal communication). Between 1969 and 1976, one of us (P. M.) installed networks of geodetic stations across the Ethiopian rift, including one near Addis Ababa⁴. Inter-station distances of up to 20 km were measured with an accuracy of 0.5–1 cm using a Mark 8 geodimeter. In January 1992 we re-measured some of these lines using GPS technology with a similar accuracy. (GPS data were obtained at 15-s intervals in three 5-h sessions using dual-frequency receivers. Processed data using precise orbits provided by Y. Bock yielded horizontal and vertical repeatabilities of ± 1 cm and ± 2.5 cm, respectively.) Astonishingly, the three long lines that we recovered, spanning 46 km of the 90-km width of the northern Ethiopian rift, yielded data agreeing with the geodimeter data within their respective experimental errors. The data indicate a total widening of this part of the rift by 1 ± 2 cm in 20 years, or a net widening rate of 0.5 ± 1 mm per year. Were this rate uniform across the whole width of the rift at this latitude, the maximum widening rate would be 1.1 ± 2.2 mm per year, much lower than any previous geological estimate.

Although convergence is permitted by the data, we consider this unlikely. Several stations on the western margin of the rift were not recovered, and many lines are yet to be measured, but those that were remeasured embrace the central most active part of this sector of the rift valley. The data imply extensive strain of

10 ± 20 nanostrain per year, an unexpectedly low strain rate for this neo plate-boundary. Episodic rifting is therefore implicit in relating our results to the manifest youthfulness of the rift faulting/fissuring^{5,6}. If strain rates are uniform in time, and if tensile failure occurs near 10 μ strain, we should not expect episodes of rifting to occur more frequently than every 3,000 years in the Addis sector.

We are reminded that, in 1932, students of Alfred Wegener installed a geodetic network spanning the Krafla dike injection zone in northern Iceland, which when remeasured in 1956 revealed no deformation⁷. The students' prescience was rewarded in 1975 when 8 m of widening occurred within their network⁸. East African rift activity is slower than in

Iceland. We approach the next millennium with interest.

Laike M. Asfaw

*Geophysical Observatory,
Addis Ababa University, Ethiopia*

Roger Bilham

Michael Jackson

*CIRES, University of Colorado,
Boulder, Colorado 80309, USA*

Paul Mohr

*Department of Geological Sciences,
University of Galway, Ireland*

1. Bilham, R. *Rev. Geophys.* **29**, 1–29 (1991).
2. Mohr, P. *EOS Trans. Am. Geophys. Un.* **68**, 721–729 (1987).
3. Gibson, I. L. *Tectonophys.* **8**, 561–565 (1969).
4. Mohr, P. *Smithsonian Astrophysical Observatory Spec. Rep.* **376**, 111 (1977).
5. Gouin, P. *International Development Research Center, Ottawa*, **118e** (1979).
6. Mohr, P., Mitchell, J. G. & Reynolds, R. G. H. *Bull. Volcan.* **43**, 173–189, 1980.
7. Niemczyk, O. & Emsehermann, E. *Spalten auf Island*, 80 (Witter, Stuttgart), 1943.
8. Björnsson, A. *et al. Nature* **266**, 318–323 (1977).

Carbon stores in vegetation

SIR — Analyses of fossil pollen and plant macrofossils in sediments and peat stratigraphies show that the American southwest has undergone extensive vegetational change since the last glacial maximum (23,000–17,000 years ago) from an area mostly dominated by woody vegetation to the deserts and arid grasslands of the present¹. Using maps published by Martin and Mehringer and estimates of carbon storage (see table), we have calculated that this vegetational change resulted in a decrease in carbon in vegetation of 12.3×10^{12} kg over the 982,000-km² study area. The change in total carbon storage of this area, however, was virtually nil, due to the accumulation of large amounts of carbonate

in arid soils (see table). The shift in vegetative cover from forest to desert simply changed the form of carbon storage — from organic carbon stored in biomass to carbonate-carbon stored in soils.

Although deserts have expanded in the American southwest since the last glacial maximum, on most other continents the area of desert land has contracted. Adams *et al.*² reported that there has been a global decrease in the area of deserts and, therefore, an increase in carbon storage in vegetation and soils since the last glacial maximum. But a comparison of their data for soil carbon storage in temperate desert (3 kg m⁻²) and semi-desert and dry steppe (6

AREAS AND CARBON STORAGE FOR VEGETATION TYPES

		Desert scrub	Desert grassland	Chaparral	Woodland	Parkland	Forest	Subalpine-alpine	Total
Area ($\times 10^{11}$ m ²) ¹	LGM	0.36	0	0.87	2.17	4.84	1.40	0.18	9.82
	Present	3.62	3.23	0	2.10	0.70	0.16	0.01	9.82
	Change	+3.26	+3.23	-0.87	-0.07	-4.14	-1.24	-0.17	0
Carbon storage (kg C m ⁻²)	Vegetation ^{4,5}	0.3	0.3	2.3	1.5	15.0	46.0	21.0	—
	Soil-organic ³	2.2	7.3	5.2	7.6	7.6	7.6	7.6	—
	Soil-carbonate ³	31.0	22.3	0.3	5.0	5.0	5.0	5.0	—
	Total	33.5	29.9	7.8	14.1	27.6	58.6	33.6	—
Carbon storage, vegetation only, entire study area ($\times 10^{12}$ kg C)	LGM	0.01	0	0.20	0.33	7.26	6.44	0.38	14.62
	Present	0.11	0.10	0	0.32	1.05	0.74	0.02	2.34
	Change	+0.10	+0.10	-0.20	-0.01	-6.21	-5.70	-0.36	-12.28
Carbon storage, vegetation+soil, entire study area ($\times 10^{12}$ kg C)	LGM	1.21	0	0.68	3.06	13.36	8.20	0.61	27.12
	Present	12.13	9.66	0	2.96	1.93	0.94	0.03	27.65
	Change	+10.92	+9.66	-0.68	-0.10	-11.43	-7.26	-0.58	+0.53
Change (%)	LGM	1.21	0	0.68	3.06	13.36	8.20	0.61	27.12
	Present	12.13	9.66	0	2.96	1.93	0.94	0.03	27.65
Change (%)	LGM	1.21	0	0.68	3.06	13.36	8.20	0.61	27.12
	Present	12.13	9.66	0	2.96	1.93	0.94	0.03	27.65

LGM, last glacial maximum.

kg m⁻²) with those of W. H. S. (ref. 3; see table) suggests that they did not include changes in soil carbonate-carbon in their assessment. Loss of desert land may be accompanied by the degradation and loss of soil carbonate with the carbon transferred to inorganic pools in the atmosphere or the oceans. As the carbonate pool is 10 times larger than the pool of soil organic carbon in deserts, the actual global increase in carbon storage due to the decrease in desert area may be significantly less than the figure reported by Adams *et al.*². Further study is needed to determine the significance of this discrepancy.

Ruth Ann Kern

Department of Botany,
Duke University,
Durham, North Carolina 27706, USA
William H. Schlesinger
Departments of Botany and Geology,
Duke University,
Durham, North Carolina 27706, USA

1. Martin, P. S. & Mehringer, P. J. Jr in *The Quaternary of the United States* (eds Wright, H. E. & Frey, D. G.) 433–451 (Princeton University Press, 1965).
2. Adams, J. M., Faure, H., Faure-Denard, L., McGlade, J. M. & Woodward, F. I. *Nature* **348**, 711–714 (1990).
3. Schlesinger, W. H. *Soil Sci.* **133**, 247–255 (1982).
4. Whittaker, R. H. & Niering, W. A., *Ecology* **56**, 771–790 (1975).
5. Ehleringer, J. & Mooney, H. A. in *Encyclopedia of Plant Physiology*, New Ser. **12D**, 205–231 (1983).

Healthy broccoli?

SIR — The interpretation by Culliton¹ of data by Talalay *et al.*² is incomplete. In fact, President George Bush's reluctance to eat broccoli may have a sound scientific basis. Despite the benefit derived from the consumption of nutrients containing inducers of enzymes that detoxify carcinogens (phase II enzymes), such as sulphoraphane, the opposite side of the equation, the formation of active carcinogenic metabolites, cannot be overlooked.

Without commenting on the wisdom of the president's lack of taste for broccoli, we would like to point out the conceptual error that phase II enzymes are detoxifying only and that the monofunctional inducers sulphoraphane, cinnamates, coumarins, thiocarbamates, diphenols and so on increase the expression of such enzymes and can thus protect against carcinogens.

It is well established that several phase II enzymes, such as epoxide hydrolase, UDP-glucuronyl transferase and glutathione *S*-transferase, are responsible for the bioactivation of many carcinogens³. Therefore, each phase II enzyme system may be considered as an activating system towards specific chemical classes, for example aromatic polycyclic hydrocarbons in the case of epoxide hydrolase or halogenated hydrocarbons in the case of glutathione *S*-transferase⁴.

There is no doubt that a magnification of phase II enzyme activities may have, simultaneously, a dual effect. If on the one hand it is possible to have an increase in detoxification and excretion for some chemicals, then on the other the many active precarcinogens point to the importance of this aspect for human health. Each post-oxidative enzyme should thus be considered as 'detoxi-toxicant' depending on the chemical involved. Taking into account that humans are exposed to a myriad of compounds, attempts to induce phase II enzymes by dietary components to reduce risk for mutations and cancer should be reconsidered carefully³.

Finally, the enthusiastic manner in which the media recommend eating broccoli in the name of a long and healthy life should be restrained by the knowledge that dietary broccoli is also able to increase several cytochrome P-450 isoforms⁵ (phase I, typical activating enzymes). Indole carbinol, a strong promoter of carcinogenesis⁶ and of inducers of dioxin-metabolizing (P450IA2) enzymes⁷, is present in large amounts in broccoli and in other members of the *Brassica* family.

Moreno Paolini

Dipartimento di Farmacologia,
Sezione di Biochimica Tossicologica,
Università degli Studi,
Via Irnerio 48, 40126 Bologna, Italy

Marvin S. Legator

Department of Preventive Medicine and
Community Health,
Division of Environmental Toxicology,
University of Texas Medical Branch,
Galveston, Texas 77555, USA

1. Culliton, B. J. *Nature* **356**, 377 (1992).
2. Talalay, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 2394–2403 (1992).
3. Paolini, M. *et al. Biochem. Pharmacol.* **42**, 457–458 (1991).
4. Paolini, M. *et al. Biochem. Pharmacol.* **38**, 2223–2225 (1989).
5. Vang, O. *et al. Chem. Biol. Interact.* **78**, 85–96 (1991).
6. Bailey, G. S. *et al. J. natn. Cancer Inst.* **78**, 931–934 (1987).
7. Bradfield, C. A. *et al. J. Tox. envir. Hlth* **21**, 311–323 (1987).

Curved saplings at Mt St Helens

SIR — Bursik and Woods¹ noted downslope curvature among standing Douglas fir and Sitka spruce saplings in forests blown down by the lateral blast of Mount St Helens on 18 May 1980. They proposed that the curvature was caused by greater shrinkage of compression wood (downslope) than normal wood (upslope) during the eruption. They concluded that sapling curvature can be used to reconstruct the temperature and duration of a lateral blast.

I question these interpretations for



FIG. 1 Living upslope-curved Pacific silver fir saplings in forest northeast of Mount St Helens 2 km from 1980 blast zone (1989 photo; 46°20.85' N, 122°01.25' W, 1,160 m). Note downslope-curved trees in background potentially caused by snow creep. Shovel handle is 0.5 m long.

two reasons. First, Bursik and Woods did not exclude other explanations, such as snow creep, for the downslope sapling curvature. Spring snowpacks at Spirit Lake commonly achieve depths of 2 m (ref. 2) and could certainly creep and bend saplings downslope. Second, if saplings northeast of the volcano are truly acting as 'bimetallic strips', an unusual amount of shrinkage-prone¹ compression wood in local tree stems, rather than the heat of the 1980 blast, may explain downslope curvature after death.

My evidence for the latter inference consists of two photographs (Figs 1, 2) showing that curvature is a natural feature of trees in forests northeast of Mount St Helens. Such stems contain abnormal quantities of compression wood in their lower 2 m (ref. 3).

Although the cause of upslope curvature in tree stems is a matter for debate⁴, such curvature in 300–500-year-old local Douglas fir (Fig. 2), and its absence in Douglas fir more than 510 years old at the same sites⁵, suggest that soil creep is responsible; the younger trees are rooted on a thick (>1 m), unstable AD 1480 tephra layer⁶. Saplings that started growing before 1980 are probably now similarly responding to creep of the AD 1800 tephra⁵ blanketing the landscape northeast of the volcano. Similar upslope curvature occurs among trees on loose volcanic debris at Lassen Volcanic National Park, California⁷.

Thus, the downslope-curved saplings noted by Bursik and Woods may be unrelated to the 1980 blast. The curvature could be related to snow creep or to warping of stems containing atypical levels of compression wood. These alternative hypotheses can be tested by studying saplings in forests outside the blast zone.

Finally, the saplings examined by Bur-



FIG. 2 Living upslope-curved 450-year-old Douglas fir at the same site. Tree diameter about 1 m (1989 photo).

sik and Woods were not Douglas fir and Sitka spruce but Pacific silver fir and western or mountain hemlock⁸. Shade-intolerant Douglas fir does not inhabit regional forest understories⁸, and Sitka spruce does not grow in inland Mount St Helens forests⁹.

David K. Yamaguchi

*Institute of Arctic and Alpine Research,
University of Colorado,
Boulder, Colorado 80309, USA*

1. Bursik, M. I. & Woods, A. W. *Nature* **355**, 594 (1992).
2. US Weather Bureau. *Climatology of the United States* Clim. Summ. No. 20-45 (Washington, DC, 1964).
3. Keimholz, R. *Am J. Bot.* **17**, 739-764 (1930).
4. Shroder, J. F. Jr. *Prog. phys. Geogr.* **4**, 161-188 (1980).
5. Yamaguchi, D. K. *Quat. Res.* **20**, 246-250 (1983).
6. Yamaguchi, D. K. thesis (Univ. Washington, Seattle, 1986).
7. Sharpe, C. F. S. *Landslides and Related Phenomena* (Pageant, New Jersey, 1960).
8. Franklin, J. F., MacMahon, J. A., Swanson, F. J. & Sedell, J. R. *Natn. geogr. Res.* **1**, 198-216 (1985).
9. Hitchcock, C. L. & Cronquist, A. *Flora of the Pacific Northwest* (University of Washington Press, Seattle, 1973).

SIR — During a recent visit to the former coniferous forest overrun by the hot pyroclastic current (blast) from Mount St Helens on 18 May 1980, Bursik and Woods¹ noted that understory saplings were not levelled like mature trees but remained upright albeit killed. They inferred that a "consistent" pattern of downslope curvature of the dead saplings occurred by differential dehydration of compression (downslope) wood relative to regular (upslope) wood — dehydration caused by the sustained hot current. They concluded that saplings therefore cannot be used as flow-direction indicators but could be used as a proxy measure of temperature and duration of a hot particulate flow. Both

conclusions are erroneous, for the premises that the curvature is consistently downslope and that the hot current directly caused the curvature are both false.

The condition of the saplings has gradually changed since the eruption. During the first 4 months after the eruption many observations were made on the timber², though mainly on large mature logs and stumps. The saplings either stood straight or had been tilted in the downcurrent direction. On volcano-facing slopes the bent saplings were concave upslope; on slopes away from the volcano they were bent concave downslope. This pattern persisted for some time, at least a few years. This eruption-induced tilting was independent of concave-upslope curvature within the basal 2 m which commonly occurs in saplings on steep slopes because of creep of regolith or winter snow (see the letter above from Yamaguchi).

While preparing a field guidebook³ in 1988, I noted a curiosity of the standing saplings that had not been present in the early 1980s: many of the dead straight or concave-downcurrent saplings had become distinctly concave upcurrent. On a 10-20° east-sloping ridge just east of Spirit Lake, where the pyroclastic current running downslope in 1980 had tilted saplings downslope, by 1988 many of them had acquired a concave curvature facing upslope (upcurrent). This curvature was mostly in the upper half of sapling trunks, unlike that near the base caused by snow creep. I noted the volcano-facing concavity because it was bizarre, fairly systematic, and different from the pattern in earlier years.

The upslope concavity in the saplings on a slope facing away from the volcano is the opposite of the downslope concavity noted by Bursik and Woods apparently on a volcano-facing slope. The volcano-facing concavity of dead saplings also exists on gentle slopes. Thus the relation of curvature to a tree's downslope compression wood is inconsistent, but its relation to blast-current direction is consistent.

The pyroclastic current of 18 May 1980 had little or no direct effect on the eventual curvature. The hot current only abraded and baked off the bark and thus exposed the sapwood on the upcurrent side, whereas the downcurrent side lay protected for one or a few years beneath an attached bark strip. By 1988 much of the bark had fallen off and been scattered by wind.

The direct cause of curvature apparently was gradual drying over a few years by differential exposure of sapwood to the ordinary perils of sun and rain. Unprotected by bark, the exposed upcurrent wood shrank, while downcurrent wood remained more or less un-

affected beneath bark, a process also seen on houses sided with unfinished wooden shingles or clapboards. Shingles or clapboards on the south side of a Northern-Hemisphere house directly exposed to the sun can warp, crack and shrink in a few years. Meanwhile the same materials on the north side of a house or where shaded by a porch roof remain pristine.

The bent saplings are reliable indicators of current direction: after a few years of weathering their acquired concavity points to the side that lost bark (upcurrent). Conversely, they are nearly useless as indicators of the temperature or duration of the pyroclastic current that killed them.

Richard B. Walt

*US Geological Survey,
Cascades Volcano Observatory,
5400 MacArthur Blvd,
Vancouver,
Washington 98661, USA*

1. Bursik, M. I. & Woods, A. W. *Nature* **355**, 594 (1992).
2. Walt, R. B. *US geol. Surv. prof. Pap.* **1250**, 439-458 (1981).
3. Walt, R. B. et al. *New Mexico Bur. Mines Min. Res. Mem.* **47**, 51-89 (1989).

Which theory?

SIR — I should like to comment on John Maddox's article¹ on the alleged elimination of divergences in quantum electrodynamics (QED) by R. Suzuki². Suzuki claims to eliminate the divergences of Feynman loop diagrams by applying to QED a transformation that decomposes the basic electromagnetic current-photon interaction into two parts, each of which makes equal and opposite contributions to the loop divergences. However, the transformation in question consists of two parts, one ordinary time reversal, the other a 'light-cone rotation' of the time-like and longitudinal photon fields. (This rotation is not consistently extended to the space-time co-ordinates or the electron fields. Its physical significance, as Maddox points out, is obscure.)

QED is invariant under the former transformation, but not under the latter. Hence, Suzuki has actually demonstrated the cancellation of divergences for a different theory, not QED. Inasmuch as the divergences of QED are invariant under that theory's symmetries — Lorentz symmetry and local gauge invariance — no transformation that leaves QED invariant can eliminate the divergences.

Maddox's article also perpetuates the myth that the renormalization of divergences is a flaw of quantum field theory. In fact, the divergences have a definite physical meaning, associated with well-understood short-distance behaviour. As long as field theories have free para-

meters, divergences pose no problem of principle. Even theories with non-normalizable divergences, suitably restricted, are perfectly sensible³⁻⁵. The demand that quantum field theories have no divergences is misguided and physically unnecessary.

Dallas C. Kennedy

*Department of Theoretical Physics,
Fermi National Accelerator Laboratory,
Mail Stop 106, PO Box 500,
Batavia, Illinois 60510, USA*

1. Maddox, J. *Nature* **353**, 497 (1991).
2. Suzuki, R. *Nuovo Cim.* **104A**, 1115 (1991).
3. Wilson, K. *Scient. Am.* **241** (2), 158 (1979).
4. Weinberg, S. in *Asymptotic Realms of Physics* (eds Guth, A. H. et al.) 1 (Cambridge, MIT, 1983).
5. Lepage, G. P. in *TASI 1989* (eds Degrand, T. & Toussaint, D.) 483 (World Scientific, Singapore, 1990).

Uncorrelated DNA walks

SIR — Peng *et al.*¹ represent DNA sequences as random walks on which a pyrimidine represents a step up and a purine a step down. On the basis of their analysis of such walks they report "long-range correlations" in the nucleotide sequences of intron-containing genes and suggest that these reveal fractal properties of genome organization which may be connected with non-equilibrium dynamic processes. This could well be true. Here I put forward a different framework for thinking about their observations which readily suggests explanations in terms of molecular biology.

Consider a random walk which is uncorrelated, that is, the direction of each step, up or down, is independent of the direction of previous steps. The walk may or may not be biased. If the probability of stepping up is the same everywhere then the root mean square fluctuation function of the walk, $F(l)$, has a slope of 0.5 on a double-log plot (this slope is the fundamental statistic of the analysis). However, if the probability of stepping up differs in different regions then, simulation shows, the slope is greater than 0.5 and can, indeed, be much greater. So the observations of Peng *et al.* of slopes greater than 0.5 in their DNA walks do not necessarily imply long-range correlations in the sense of a lack of independence at long ranges. They may instead reveal that, for example, introns and exons are statistically different with a biased walk in the exons

and an unbiased walk in the introns. That the walk is biased in the exons is strongly suggested by Peng *et al.*'s descriptions of the cDNA walks. Furthermore, if we calculate the probability of stepping up at the level of the codon, using the amino-acid frequency data in ref. 2 and the codon usage data in ref. 3, we find that the probability of a step up is 0.46 in GC-rich regions of the human genome and 0.48 in AT-rich regions.

What these latter data do not reveal is the origin of the bias reversals which seem present in Peng *et al.*'s cDNA. The global min-max partitioning procedure used by Peng *et al.* to analyse their walks will restore slopes of 0.5 as long as there are no more than three statistically distinct regions on the walk. That this procedure 'works' for the cDNA means that there are not many reversals of bias in the coding regions of genes. This suggests that the ultimate source of the bias reversals may be the mosaic structure of genomes with regard to GC content, which is more complex than a GC-rich/AT-rich dichotomy (ref. 4).

Sean Nee

*AFRC Unit of Ecology and Behaviour,
Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK*

1. Peng, C.-K. *et al.* *Nature* **356**, 168–170 (1992).
2. Jukes, T. H., Holmquist, R. & Moise, H. *Science* **189**, 50–51 (1975).
3. Sharp, P. M. *et al.* *Nucleic Acids Res.* **16**, 8207–8211 (1988).
4. Bernardi, G. *et al.* *Science* **228**, 953–958 (1985).

Selfish genes in mosquitoes

SIR — Hurst and colleagues^{1,2} state that "within any population of [the mosquito] *Culex pipiens* there are two sorts of individual, those that bear/harbour . . . *Wolbachia* [bacteria] and those that do not". But, according to Yen and Barr³, all wild-type *C. pipiens* appropriately examined before 1973 were found to carry *Wolbachia*. It was only when Yen and Barr produced uninfected individuals artificially by tetracycline treatment that it was found that matings of uninfected females × infected males were sterile, whereas all matings by uninfected males were fertile.

Laven⁴ test-crossed allopatric *C. pipiens* strains and found unidirectional or bidirectional incompatibility (sterility) in many cases. The crossing types are cytoplasmically inherited and the relationships between various strains are too complex to explain by a simple 'presence-absence' hypothesis. The same applies to variability in compatibility type found within wild populations^{5,6}. It is presumed that different strains have

genetically different *Wolbachia* whose interactions somehow cause the observed incompatibility^{2,3}. The presumed differences in the *Wolbachia* between *C. pipiens* strains should be detectable by molecular techniques, for example as described in ref. 7.

It has long been realized that a mixed population of two unidirectionally cytoplasmically incompatible types would be expected to show selection for the type whose females are not sterilized in cross matings^{2,4,8,9}. Strong selection for *Wolbachia*-infected *Drosophila* has been observed in Californian wild populations¹⁰.

Cytoplasmic incompatibility has never been reported in *Anopheles* (malaria) mosquitoes and two *Anopheles* species examined by Yen (cited in ref. 11) did not carry *Wolbachia*. The *Wolbachia* associated with cytoplasmic incompatibility in diverse insects are very closely related⁷.

It might be possible to introduce them into *Anopheles* mosquitoes by injection, or even by feeding. If the normal 'rules' about unidirectional incompatibility and maternal inheritance apply, it should be possible to initiate selective replacement of an uninfected wild-type *Anopheles* population by 'seeding' with a few of the artificially infected *Anopheles*. If a gene for inability to transmit malaria^{12,13} could be inserted into the *Wolbachia* before introducing the bacteria into the vector population, the selective replacement process should render the population harmless to man without the need for mass release of the non-transmitting strain. This method may prove more feasible than the use of a transposable element such as a *P* factor¹⁴.

C. F. Curtis

*Department of Medical
Parasitology,
London School of Hygiene & Tropical
Medicine,
University of London,
London WC1E 7HT, UK*

1. Hurst, G. D. D., Hurst, L. D. & Majerus, M. E. N. *Nature* **356**, 659–660 (1992).
2. Hurst, L. D. *J. theor. Biol.* **148**, 269–277.
3. Yen, J. & Barr, A. R. *J. Invert. Path.* **22**, 242–250 (1973).
4. Laven, H. in *Genetics of Insect Vectors of Disease* (eds Wright, J. W. & Pal, R.) 251–275 (Elsevier, Amsterdam, 1967).
5. Subbarao, S. K., Krishnamurthy, B. S., Curtis, C. F., Adak, T. & Chandras, R. K. *Genetics* **87**, 381–390 (1977).
6. Barr, A. R. *Nature* **283**, 71–72 (1980).
7. O'Neill, S. L., Giordano, R., Colbert, A. M. E., Karr, T. L. & Robertson, H. M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2699–2702 (1992).
8. Caspari, E. & Watson, G. S. *Evolution* **13**, 568–570 (1959).
9. Fine, P. E. M. *J. Invert. Path.* **30**, 10–18 (1978).
10. Turelli, M. & Hoffmann, A. A. *Nature* **353**, 440–442 (1991).
11. Wright, J. D. & Wang, B. T. *J. Invert. Path.* **35**, 200–208 (1980).
12. Collins, F. H. *et al.* *Science* **234**, 607–609 (1986).
13. Feldmann, A. M. & Ponnudurai, T. *Med. Vet. Ent.* **3**, 41–52 (1989).
14. Curtis, C. F. & Graves, P. M. *J. trop. Med. Hyg.* **91**, 43–48 (1988).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references.

Mad rivers run sane?

Anna Scherbakova

Ecocide in the USSR: Health and Nature Under Siege. By Murray Feshbach and Alfred Friendly Jr. Basic Books: 1992. Pp. 376. \$24. (To be published in the United Kingdom by Aurum Press later this year.)

THE official attitude of the former Soviet Union towards the environment remained remarkably constant and emphasized the need to transform nature in pursuit of a modern industrial society. The Bolshevik view of the world in terms of *kto-kogo* (who conquers whom) also applied to nature and was reflected in Lenin's conception of Communism as "Soviet power plus the electrification of the whole country", as well as in Gorky's more poetic view that the Soviet people would make "mad rivers run sane". It is the great merit of *Ecocide in the USSR* that Murray Feshbach and Alfred Friendly capture not only the magnitude and detail of environmental abuse caused by this adversarial view of nature, but also interpret it in terms of the historical, political, economic and social complexities of Soviet life.

The authors' account is a frightening one, as might be expected for a country in which even today Russian governmental ministers can seriously propose the use of "peaceful nuclear explosions" as an environmentally sound way to dispose of toxic waste, spent nuclear fuel and chemical weapons. Most shocking for Western readers is likely to be the well-documented record of the medical consequences of the Soviet Union's failure to invest adequately in health care and environmental protection. This record includes the decline in life expectancy for Soviet men between 1964 and 1989 (from 66.1 to 63.8 years), the rise in infant mortality in the 1970s and 80s (to levels comparable to those in China and Sri Lanka) and the determination in 1988 that more than half of all Soviet schoolchildren were in poor health.

Although the lack of reliable Soviet environmental data and health statistics impedes the establishment of definitive cause-and-effect relationships, the authors marshal impressive evidence linking environmental pollution and disease. They report, for example, that in the 68

Soviet cities with the greatest air pollution (regularly measured at ten times the permissible maximum), the rate of illness among the 40 million residents was more than 50 per cent higher than the national average. According to Aleksei Yablokov, an advisor to President Yeltsin and one of the best-informed environmental advocates in the former

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Air pollution in Russia — a staggering array of environmental and health problems now confronts the newly independent but economically hard-pressed states of the former Soviet Union.

Soviet Union, there are regions where, owing to the overuse of pesticides, no healthy youths of military-service age can be found and where cancer would appear to be the only cause of death.

Feshbach and Friendly are especially good at ferreting out obscure but important Russian-language sources and display an almost encyclopaedic grasp of environmental issues throughout the former Soviet Union. The book also profits from a comparative approach that relates statistics on Soviet health and environmental problems to those in the

West. It is useful if disturbing to learn, for example, that Soviet infant mortality rates are comparable to those in East Harlem and Washington DC as well as in China. Perhaps most remarkable about the finely documented study is that it remains immensely readable.

The authors find very little positive to report on environmental developments in the Soviet Union. This is not surprising given seven decades of official neglect. The one exception is the awakening of an environmental consciousness among the Soviet people and their leaders caused by the Chernobyl nuclear accident in April 1986. That consciousness, the authors suggest, may lead to policy change.

Indeed, Chernobyl has led to a more fact-conscious, socially responsible, and

probing news media. It also had a catalytic effect on Soviet popular thinking about the need for concerted action to protect the environment and resulted in the proliferation of new ecology movements, many with nationalist agendas. This surge of environmental awareness, most pronounced in Ukraine, Belarus (Byelorussia) and the Baltic States, reflected learning at the grassroots level of at least two important lessons: (1) that the Soviet people had been misled and betrayed by a government that had promised to protect and care for them; and (2) that Moscow would tolerate a surprising degree of criticism on environmental issues, which could be used to promote regional interests *vis-à-vis* the centre. Criticism of nuclear power and the government's environmental policy, in short, represented a politically safe surrogate for broader criticism of society.

Feshbach and Friendly perceptively describe the *glasnost* and nationalist fallout of Chernobyl. Less adequately addressed is the post-

independence phenomenon in which former environmental activists succumb to the same kind of economic demands that frustrated their Soviet predecessors. How sad but ironic today for leaders of Lithuania, Armenia and soon perhaps even Ukraine, to embrace as economically indispensable the same unsafe nuclear power plants that only yesterday were the targets of their environmental wrath.

It remains to be seen if the stirrings of environmental consciousness and reform after Chernobyl will blossom in the

newly independent, but economically hard-pressed states of the former Soviet Union. How much of the reported \$10 billion deal with Chevron will the Kazakh government devote to cleaning up the environment and improving public health? Will energy-starved Armenia restart the two nuclear reactors shut down in 1989 without making them any more resistant to earthquakes? Is non-Soviet Uzbekistan likely to be any more able than its Communist predecessor to free itself from a chemically ruinous and water-squandering cotton monoculture?

Feshbach and Friendly do not attempt to answer these difficult kinds of questions. Their meticulously researched book does, however, provide an invaluable chronicle of the environmental damage wrought by a system that treated natural resources as inexhaustible free goods and the vast Soviet wilderness as an adversary in the battle to fulfil the next five-year plan. □

Anna Scherbakova is at the Center for Russian and Eurasian Studies, Monterey Institute of International Studies, 425 Van Buren Street, California 93940, USA.

Tales of transformation

John Cairns

Origins of Human Cancer: A Comprehensive Review. Edited by Joan Brugge, Tom Curran, Ed Harlow and Frank McCormick. Cold Spring Harbor Laboratory Press: 1991. Pp. 204. \$80.

THE Cold Spring Harbor Laboratory has held two meetings on the origins of human cancer, one in 1976 and the other in 1990. A comparison of the two books of proceedings gives a fascinating picture of the development of cancer research over the intervening years.

By 1976, President Nixon's war against cancer was underway, and money was flowing into the search for a cure and into the study of the underlying molecular biology of carcinogenesis. For complicated reasons, there was little enthusiasm for the idea of prevention and so epidemiologists did not share in these riches; but they were the people who knew most about the origins of human cancer, and about half of the 1976 meeting was taken up with their account of the epidemiology of cancer. The rest of the meeting was mostly concerned with the biology of tumour viruses and the mechanisms of chemical carcinogenesis. This was the time that assays such as the Ames test were starting to be used, and it seemed conceivable that they would quickly pinpoint the causes of most forms of cancer, with the search for human tumour viruses identifying the other causes. It is instructive to see how many of these hopes have remained unfulfilled, even though our understanding of the molecular biology of cancer has raced ahead.

First, the disappointments. Since 1976, epidemiologists have had a difficult time; the search for the causes of breast and colon cancer is proving to be unexpectedly difficult, because there seems to be no single factor that could be linked to some plausible mechanism of carcinogenesis and could account for the

widely different cancer rates in the West and the Far East. The short-term tests for carcinogenicity have not shown us how to prevent cancer; their predictive value is too low, and it is now not even clear how many of the major cancers in industrialized nations are the result of exposure to environmental mutagens. Even when a cause is known, the knowledge may not be put to use; for example, little has been done to stop tobacco companies extending their business to developing nations, which shows that the world is still not keen on prevention. The reason may be that the war against cancer was instigated by clinicians, and the main thrust has therefore been to develop new forms of treatments. But here too, the objective is proving difficult to reach; readers of *Nature* have about the same overall chance of dying of cancer next year as their parents had when they were the same age.

Most of the 1990 meeting, however, was concerned with the successes. A few new causes of cancer have been discovered, of which hepatitis B virus and certain strains of human papilloma virus are the most important; together they may account for almost as many cancers as cigarettes. Rid the world of these three agents, and you have solved half the problem.

The transforming oncogenes of DNA and RNA tumour viruses were discussed at the 1976 meeting, but I do not think anyone present guessed how quickly this field was to develop. The turning point came in 1982, with the discovery that human tumours often show changes in genes whose counterparts had been identified in bird and rodent retroviruses. There followed an explosion of information thanks to the evolving technology for identifying and comparing nucleic-acid sequences.

About two-thirds of the 1990 meeting was therefore devoted to the function and alterations of oncogenes and tumour-suppressor genes. The functions of the products of these genes are more

or less what one might have predicted — the transduction of signals between cells and the control of cell division and differentiation — and they operate as signal receptors and transducers or as modulators of enzyme action or as regulators of gene expression. The picture that emerges is both simple and complex. Simple, because the number of such genes is not unmanageably large and certain cancers fairly regularly show alteration in particular oncogenes. Complex, because we now see that the regulation of cell behaviour is achieved by an immensely complicated set of interactions between proteins and nucleic acids, and this adds another dimension to all those already complex diagrams of intermediary metabolism. And there is the further dimension of time; particular patterns of interaction are transient and normally not repeated in the life of an organism, and that is why many of the mutations studied by developmental biologists prove to be in what are now called oncogenes.

We can surely expect that all this bustling molecular biology will soon be illuminating the origins of human cancer. In 1982, it was already possible to see the day coming when the molecular biologist would be able to look at the sequence changes in each class of cancer and tell the epidemiologist what kind of mutagens to look for in the patient's past. This has now come to pass, although so far the epidemiologists have been told only what they already know — for example, that skin cancers show the kinds of sequence changes in the p53 gene that one would expect to be produced by ultraviolet light, and that familial susceptibility to certain cancers can be associated with germline mutations in tumour-suppressor genes.

The 80 or so articles in the new *Origins of Human Cancer* give a wonderful overview of these developments. What would I not give for a preview of the discoveries of the next 14 years. □

John Cairns, formerly in the Department of Cancer Biology at the Harvard School of Public Health, Boston, USA, is at Holly Grove House, Wilcote, Chipping Norton; Oxfordshire OX7 3EA, UK.

Corrections

■ In his review of *The Cybernetics Group* (*Nature* **356**, 117; 1992), Donald Michie complained of a lack of references to the writings and activities of "the new cybernetics tradition". The references are in fact to be found in Note 21 on page 326 of the book.

■ Andrew Warwick, who reviewed *The Maxwellians* in *Nature* **357**, 291 (1992), is a fellow at St John's College, University of Cambridge, as well as being affiliated to the Department of History and Philosophy of Science, as reported.

Unlocking the past

Douglas Palmer

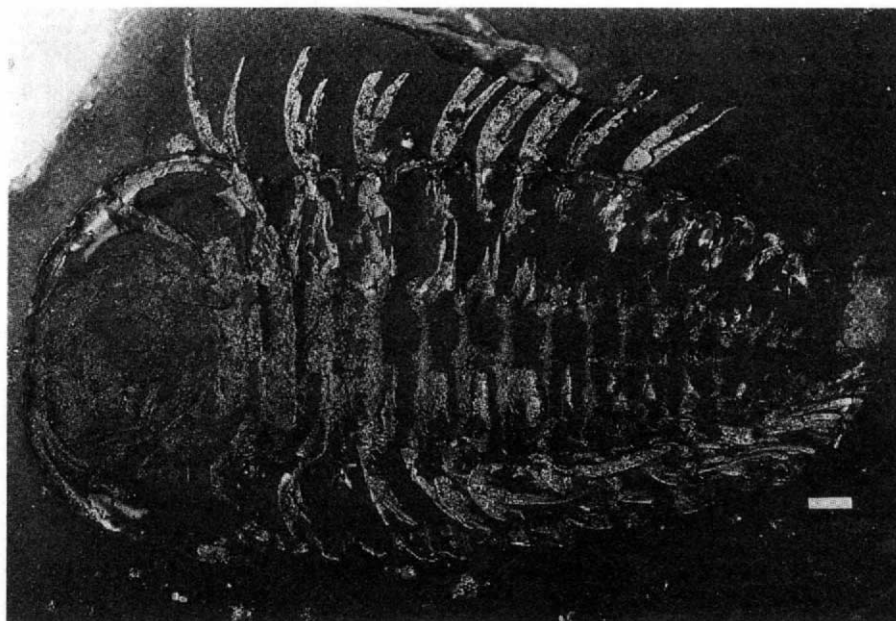
Taphonomy: Releasing the Data Locked In the Fossil Record. Edited by Peter A. Allison and Derek E. G. Briggs. Plenum: 1991. Pp. 560. \$95, £67.06.

THE classical descriptive terms of palaeontology — 'petrified fossils', 'the vestiges of past life' — all suggest loss of substance and with it information. It is all too easy to take a negative view of the great stone sarcophagus that is the fossil record. But just how representative is this record of the great variety and relative abundance of past life? And why and how are some organisms preferentially preserved? The study of burial processes, or 'taphonomy', addresses these and the related problems of how the postmortem remains of organisms are recruited to the sedimentary record.

Because of the enormous diversity of literature for multidisciplinary subjects such as taphonomy, well-written and organized reviews play an important part in helping all concerned keep in touch with the latest developments. This new volume gives a detailed and positive view of the nature of the fossil record and its potential for enriching our understanding of evolution. The editors have brought together an excellent and timely survey of taphonomy today. It builds on the earlier volumes in the series (on skeletal growth, animal-sediment relations and biotic interactions) and complements other recent works such as Joseph Carter's encyclopaedic *Skeletal Biomineralization* (Van Nostrand Reinhold, 1990) and Steven Donovan's *The Processes of Fossilization* (Belhaven, 1991).

Of particular value are the introductory discussions on organic biomolecules and nonmineralized tissues. The study of biomolecules in the fossil record is opening up new vistas. My only regret is that the chapter on this subject is not longer. Also welcome are the analyses of the role of the main mineral groups in the fossilization process. Iron sulphides, phosphates, diverse carbonates and silica minerals are all given substantial review chapters.

Most of the authors conclude by emphasizing some current problems and unanswered questions. Why is there a great fluctuation through time in the volume of silica locked in sedimentary rocks and particular fossil groups? Which came first, the change in available marine silica or the blooms and radiations of silica-fixing organisms? Why do ancient sediments and fossils display a



Ageing gracefully — exceptional preservation of a trilobite's appendages by pyrite (*Triarthrus eatoni*, Upper Ordovician, New York State, USA). Photo by Dr J. E. Almond from H. B. Whittington's *Trilobites* (Boydell, in the press). Scale bar, 1 mm.

much greater range of pyrite formation than modern sediments with their incipient fossil shells? As Donald Caulfield and Robert Raiswell say, "early diagenetic processes are unable to explain many features of the fossil pyritization". This stimulating volume should be available to all students of palaeontology. It will

be their future work that will help release some of the vast supply of unexplored data still locked away in the fossil record. □

Douglas Palmer is a freelance palaeontologist and science writer at 31 Mawson Road, Cambridge CB1 2DZ, UK.

Of compasses and currents

N. David Mermin

Icons and Symmetries. By Simon L. Altmann. Oxford University Press: 1992. Pp. 104. £14.95, \$29.95.

WRITING accurately and non-technically about technical matters is a challenging art. In these lectures on symmetry, Simon Altmann has produced three fascinating specimens of the genre. Originally given to the Faculty of Science of the Catholic University of Leuven, the lectures, in their written form, are now also addressed to young people not yet at university (in the United Kingdom) or who are in general liberal-arts college science courses (in the United States).

Altmann begins by retelling the famous tale of Ørsted, the compass and the current. Symmetry appears to dictate that the needle cannot be deflected from the north by a north-flowing current, so any electromagnetic effect should clearly be sought with the wire perpendicular to the needle, yet none is found. Appearances can be deceiving. This particular deception delayed for some time one of the great unifying discoveries of science. Altmann presents the episode as an

instance of confusing the properties of an icon (in this case an arrow on a page) and the object it is intended to represent (in this case a directed line segment or vector). This well-nigh universal use of arrows overlooks the fact that there are two very different ways a cylindrical rod can be used to point. One way puts a special mark on the end of the rod. But one end can also be distinguished from the other by spinning the rod about its axis. The two ways correspond to polar and axial vectors, which behave oppositely under reflection in a plane. Ørsted's discovery was delayed by the failure to realize that the icon of the arrow captures only one of the two types of vectors.

The story is told with considerable charm and many attractive and informative figures. We are given advice, some good, some bad, from luminaries ranging from St Thomas Aquinas to Pierre Currie on the relation one should expect between the symmetries of causes and their effects. I could happily build a week or two of a liberal-arts physics course around the text of Altmann's first lecture.

But the second or third lectures might tax even the well-educated sixth-former.

In the second lecture, Altmann teaches us about quaternions (as well as, rather briskly, complex numbers) and how Hamilton failed to carry his celebrated vision on Brougham Bridge through to its proper conclusion. If Ørsted confused an axial with a polar vector, Hamilton's mistake was to confuse it with a 180 degree rotation about that axis. As a result it was left to Rodriguez to derive with complicated spherical geometry what Hamilton ought to have found trivially with his quaternions: the relation between the product of two rotations and their half angles. If Ørsted was deluded by taking his arrow too literally, Hamilton fell under the spell of his quaternion units i , j and k , whose squares were -1 , like the square of the complex unit i . He thus could not resist associating them with 90 degree rotations, thereby missing their true character as rotations through 180 degrees.

In the final and most ambitious lecture, Altmann discusses the instability of a one-dimensional conducting crystal against spontaneously deforming into an insulating crystal with a larger unit cell. The prediction of this now celebrated effect existed for many years only as a parenthetical aside in Rudolf Peierls's wonderful textbook on solid-state physics. To describe the Peierls transition to his general readers, Altmann develops in a few pages the concepts of energy levels, degeneracies, orbitals, Brillouin zones and much of the other apparatus of the independent electron theory of metals, even proving that a glide plane implies a twofold degeneracy at the zone boundary. It is a heroic and highly original effort, which I shall certainly return to for inspiration should I ever feel courageous enough to make such an attempt myself. But undergraduates in the United States, sixth-formers and even mature scientists who have not survived a course in quantum mechanics may well find it puzzling.

In launching this pedagogical *tour de force*, Altmann remarks that "it is very easy to talk rubbish when discussing degeneracy and, even if one is not aiming at much rigour, a minimum of decency in the treatment is absolutely essential if one is not to fall into nonsense." Anyone presenting serious science to the nonspecialist would be well advised to take to heart the general principle enunciated here. Altmann has made a first-class effort in these essays, rarely sacrificing accuracy for the illusion of intelligibility, always writing with good humour and grace, and maintaining considerable decency in the treatment. □

N. David Mermin is in the Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501, USA.

Source book

Andrew Miller

Macromolecular Crystallography with Synchrotron Radiation. By John R. Helliwell. Cambridge University Press: 1992. Pp. 595. £95, \$165.

'HAVE crystals, will travel.' This is a handbook for gipsy crystallographers seeking the brightest X-ray beams. Most of what we know about the molecular machinery of life comes from X-ray crystallography. When a biological macromolecule (enzyme, antibody, transfer RNA) or a macromolecular complex (virus, ribosome) can be crystallized, there is the potential, by X-ray crystallography, to determine the positions of all its atoms and hence to understand how it functions. The scattering power of these crystals is low and the crystals themselves are often small, so the brightest X-ray beams are required. The European Synchrotron Radiation Facility, which begins operating in 1993 and becomes available to users in 1994, will provide the brightest continuous X-ray beam in the world. John Helliwell has now written a valuable comprehensive survey of this research field.

Helliwell has adopted a catch-all approach to his wide-ranging subject. There are sections on methods of crystallization, elementary principles of macromolecular structure, basic crystal symmetry, the elements of X-ray diffraction, details of specific methods such as the Laue technique and multiple wavelength anomalous diffraction, time-resolution and molecular dynamics. There are chapters on the nature of synchrotron radiation, on storage rings, and on the range of sources that can now be designed, such as insertion devices, wigglers, wavelength shifters and tapered, tunable undulators. Beamlines receive X rays from a specific source and then condition the beam for use, and the beamline components, such as mirrors, monochromators and detectors, are also described. The subjects of many of these chapters could, and often do, have books to themselves. Helliwell has provided a valuable service by gathering such a diversity of material in one volume. The book will be a boon to final-year research students completing their theses (although few will be able to afford it).

Adequate theoretical foundations are provided throughout, but this is really an experimentalist's book, with emphasis on the practical details of working at synchrotron radiation sources. Part of the book is a sort of consumer's guide to synchrotron sources of the world, with lists of the beamlines available and their

physical parameters. The practical details, which all too frequently are impossible to find, are particularly helpful to a newcomer to synchrotron radiation, and the assured tone of these sections is due to the fact that Helliwell has been involved in many experiments, justifying what might otherwise seem to be a surfeit of self-references. There are excerpts from synchrotron radiation laboratory manuals and helpful tables.

The book is written in a clear and simple, if somewhat breathless, style, with jargon defined and explained. It does not have the intellectual coherence of works by A. Guinier (*X-ray Diffraction*, Freeman) or W. Cochran and H. Lipson (*The Crystalline State*, Vol. 3, Bell) — the field is too new for that. And advances in synchrotron radiation research are occurring so rapidly that even this book, which has references to 1991 publications, is already out of date, but not damagingly so and certainly not through any fault of the author. Helliwell occasionally overreaches himself in his effort to be fully comprehensive; for example, there is only about half a page on the complex subject of inelastic neutron scattering in the investigation of molecular dynamics.

The unsung heroes of many of the spectacular results coming from synchrotron radiation are the scientific staff at the facilities who often show imagination and originality in building instruments that cope with extremes of radiation and that still perform superbly. Protein crystallographers rely on these innovators as much as on the biochemists who provide the crystals, and Helliwell's book gives them the full recognition they deserve. □

Andrew Miller is in the Department of Biochemistry, University of Edinburgh, Edinburgh EH8 9XD, UK.

New in paperback

■ *Prime Paradigms: Sex Roles and Social Bonds* by Linda Marie Fedigan, now with a new introduction. "Courageous and hard-hitting" was how John H. Cook described this work in his review in *Nature* **302**, 359 (1983). University of Chicago Press, \$21.75, £15.25.

■ *Changing Order: Replication and Induction in Scientific Practice* by H. M. Collins, now with a new afterword. Reviewed in *Nature* **319**, 274 (1986).

■ *The Holy Land: An Archaeological Guide from Earliest Times to 1700*, 3rd Edn by Jerome Murphy-O'Connor. Reviewed in *Nature* **288**, 503 (1980). Oxford University Press, £8.99.

■ *The Dark Romance of Dian Fossey* by Harold Hayes, a compelling biography of the woman who abandoned her career as an occupational therapist to live in Africa among the last remaining mountain gorillas. Corgi Books, £6.99.

Human gene therapy comes of age

A. Dusty Miller

Advances in the understanding of molecular biology of human disease and the development of efficient gene transfer techniques have resulted in practical approaches to human gene therapy, with new techniques being developed at an increasing rate. The first trials have now begun in humans and initial results are positive.

THE treatment of human disease by gene transfer has now moved from the theoretical to the practical realm. The first human gene therapy trial was begun in September 1990 and involved transfer of the adenosine deaminase (ADA) gene into lymphocytes of a patient having an otherwise lethal defect in this enzyme, which produces immune deficiency. The results of this initial trial have been very encouraging, and have helped to stimulate further clinical trials.

Many assumptions about the nature of initial human gene therapy trials have proved inaccurate. The first approved trial of gene transfer into humans in May 1989 did not involve gene therapy at all, but was performed to track tumour-infiltrating lymphocytes after infusion into patients with melanoma. Many had assumed that haematopoietic stem cells would be the first target of gene therapy because of the long-term persistence of these cells and the variety of diseases that affect the many cell lineages that are derived from stem cells. But difficulties with efficient gene transfer and appropriate gene expression in this system have slowed this approach. Instead, work has focused on more differentiated somatic cells, as evidenced by an approved human gene therapy trial involving the transfer of low-density lipoprotein (LDL) receptor to hepatocytes. Finally, gene therapy may have more immediate application to acquired diseases, such as cancer or infectious disease, rather than the single-gene genetic defects originally proposed as primary targets. An example is a clinical trial of the use of cytokine gene transfer into tumour cells to stimulate host immune response against these tumour cells.

Here I will summarize the present status of human gene therapy trials, describe exciting new results that may soon lead to human trials, and attempt to identify areas of research that will be important for the evolution of human gene therapy. In part, this review has grown out of a meeting at the National Institutes of Health in Bethesda in December 1991, where the explosion of new ideas and major advances was very apparent.

Human gene therapy approval process

Proposals for human gene therapy in the United States must pass several levels of review. There is much overlap in the purview of the numerous committees, but given the novelty of human gene therapy, this redundancy is justified. The principal issues that are addressed are the safety of the procedure for both the patient and the general public, the anticipated benefit to the patient in comparison to the potential risk, and in the case of marking experiments where the patient probably will not benefit directly from the procedure, the likelihood that useful information will be derived from the clinical trial. These issues are detailed in the guidelines for human gene transfer¹ adopted by the Recombinant DNA Advisory Committee (RAC). The Food and Drug Administration (FDA) has adopted its own set of guidelines² which primarily deal with the characteristics, production and certification of the biological materials used for gene transfer. Table 1 lists the human gene transfer protocols that have been approved by the RAC and all of its subcommittees. Some of these protocols have not received final approval

from the FDA, but this should be regarded as more of a technical than a conceptual hurdle.

Approved human gene marking trials

All of the cell-marking protocols approved so far (Table 1) use retroviral vectors to transfer marker genes into cultured human cells that will be returned to the same patient from whom the cells were obtained. The first class of marking experiments involves the use of the marker to follow cells having presumed therapeutic use in patients. One such cell type, tumour-infiltrating lymphocytes, can be recovered from tumours, grown *in vitro*, and infused in large numbers into patients. This procedure results in clear clinical improvements in many cancer patients, although in most cases the disease recurs³. Cell marking is used to study the persistence and homing to tumour sites after reinfusion. Retroviral vectors have proved safe and useful in initial experiments⁴, although the information gained from these experiments has not yet led to improvements in tumour-infiltrating lymphocyte therapy.

Cell marking will be used to follow donor hepatocytes used to treat acute hepatic failure. Hepatocytes will be marked with retroviral vectors in short-term culture because hepatocytes are difficult to grow and may not engraft after prolonged culture. Because of this, it is difficult to select for the presence of the transferred drug marker, thus the number of hepatocytes that are marked is variable. But initial experiments⁴⁴ have shown that a significant number of the hepatocytes can be marked and, with polymerase chain reaction (PCR)-based techniques for detection of vector DNA, the modified hepatocytes should be detectable in liver biopsies. Such marking will allow the effectiveness of hepatocyte transplantation to be assessed.

One of the most recently approved cell marking protocols involves marking human immunodeficiency virus (HIV) antigen-specific T cells that can be grown in culture and that will be used in patients to kill HIV-infected cells. The cells will be used in the context of bone marrow transplantation for treatment of lymphoma, a common complication of AIDS. AIDS patients with non-Hodgkin's lymphoma will be given bone marrow from normal matched donors after high-dose chemotherapy and total body irradiation to eradicate lymphoma, which also destroys the patients' bone marrow cells. Normal marrow from matched donors will then be administered to replace the patients' haematolymphoid system. In an attempt to protect the incoming normal marrow from being infected by residual HIV in the patient, Zidovudine (AZT) will be administered to block virus spread by inhibiting viral reverse transcription, and HIV antigen-specific killer T cells will be administered to kill residual HIV-infected cells.

As an interesting safety feature, the marking vector in this study confers both hygromycin resistance and ganciclovir sensitivity owing to the expression of a bifunctional protein encoded by portions of the bacterial hygromycin phosphotransferase and herpes simplex virus thymidine kinase genes⁵. Thus vector-transduced cells can be positively selected in culture to ensure that the infused cells all carry the vector, and can be negatively selected in the patient by administration of ganciclovir. This

TABLE 1 Approved human gene transfer trials

Experiment type	Description	Principal investigator	US Institution
Cell marking	<i>neo</i> marker transfer to tumour-infiltrating lymphocytes	S. A. Rosenberg	National Institutes of Health (NIH)
		M. T. Lotze	University of Pittsburgh
		J. S. Economu	University of California, Los Angeles
	<i>neo</i> marker transfer to hepatocytes used to treat acute hepatic failure	F. D. Ledley	Baylor College of Medicine, Houston
	<i>hph-tk</i> marker transfer to HIV antigen-specific T cells used for treatment of AIDS	P. D. Greenberg	Fred Hutchinson Cancer Research Center, University of Washington
	<i>neo</i> marker transfer to leukaemic cells that may be present in autologous marrow used for treatment of leukaemia	M. K. Brenner	St Jude Children's Research Hospital, Memphis
		A. B. Deisseroth	M.D. Anderson Cancer Center, University of Texas
		K. Cornetta	Indiana University, Indianapolis
	<i>neo</i> marker transfer to neuroblastoma cells that may be present in autologous marrow used for treatment of neuroblastoma	M. K. Brenner	St Jude Children's Research Hospital, Memphis
Gene therapy	ADA gene transfer to lymphocytes from ADA-deficient patients	R. M. Blaese	NIH
	LDL receptor gene transfer to hepatocytes from LDL receptor-deficient patients	J. M. Wilson	University of Michigan
	Tumour necrosis factor gene transfer to tumour-infiltrating lymphocytes	S. A. Rosenberg	NIH
	Tumour necrosis factor gene transfer to tumour cells	S. A. Rosenberg	NIH
	IL-2 gene transfer to tumour cells	S. A. Rosenberg	NIH
	Antigen transfer to tumour cells <i>in situ</i> to stimulate immune rejection	G. J. Nabel	University of Michigan, Ann Arbor
	Toxin gene transfer to ovarian carcinoma cells for treatment of ovarian cancer	S. M. Freeman	University of Rochester
	ADA gene transfer to peripheral blood stem cells from ADA-deficient patients	R. M. Blaese	NIH

These trials have been approved by the Recombinant DNA Committee of the NIH as of February 12, 1992, but have not necessarily been finally approved by the US Food and Drug Administration. Applications for these trials are published regularly in *Hum. Gene Ther.*

would allow selective killing of the transduced T cells in patients if a problem arises with the cells, such as uncontrolled growth, or with too vigorous a response of the T cells against HIV-infected cells.

The second class of cell-marking protocols (Table 1) involves an attempt to detect residual cancerous (leukaemia or neuroblastoma) cells in marrow infused into patients during autologous transplantation for disease treatment. In these studies, bone marrow is collected from a patient before aggressive chemotherapy and radiotherapy for ablation of cancerous cells in the patient. The bone marrow is then reinfused to rescue the patient from this otherwise lethal ablative therapy. Relapse of disease in a patient may be due to residual cancerous cells in the patient or in the infused marrow. The gene-marking experiment is an attempt to resolve this question and, based on the results, to develop improved strategies for ablation of cancerous cells in the patients or for removal of these cells from the infused marrow. The difficulty in these experiments is that a very low number of cancerous cells in the infused marrow might cause recurrent disease and current techniques mark only a fraction (about 1 to 10%) of the cancerous cells present. Thus the ability to resolve the question of the source of recurrent disease is limited.

Approved gene therapy trials

All of the gene therapy trials approved so far (Table 1) use retroviral vectors for gene transfer into cultured human cells that will then be administered to patients, with the exception of a protocol that employs liposomes to transfer antigen-

encoding plasmid DNA directly into tumour cells *in situ*. The first class of experiments involves the transfer of genes to treat single-gene inherited disorders, the classical targets envisioned for gene therapy. The first human gene therapy trial involves treatment of ADA deficiency by transfer of the ADA gene into patients' T-cells, the cells that are most affected by this disorder. ADA deficiency leads to high levels of 2'-deoxyadenosine in the circulation which is toxic to both T and B cells and results in severe combined immunodeficiency. ADA deficiency is a lethal disorder that can be corrected by bone marrow transplantation, but unfortunately only about one third of such patients have suitably matched donors. The disease can also be partially corrected by weekly injections of bovine ADA that has been conjugated to polyethylene glycol (PEG-ADA) to increase its half-life in the circulation. Although PEG-ADA treatment can reduce systemic 2'-deoxyadenosine levels and result in a marked disease improvement, it is becoming clear, in part from the ongoing gene therapy trial, that it is important to produce ADA within T cells for more complete correction of the disease phenotype.

In the current gene therapy trial, two ADA-deficient patients receiving PEG-ADA are also being treated by infusions of their own cells that have been transduced with a retroviral vector that expresses human ADA and neomycin as a selectable marker. The first patient has received eight infusions over 10 months. Significantly, the ADA level in circulating cells has increased from <2% to 20% of normal and the circulating lymphocyte count has gone from below normal ($570 \mu\text{l}^{-1}$) with only PEG-ADA therapy into the normal range ($2,100 \mu\text{l}^{-1}$). Immune func-

tion in both patients has improved significantly as measured by isohaemagglutinin titres, skin tests for antigen sensitivity, and cytotoxic T cell assays. Contrary to early fears that the therapy would have no long-term benefit because of a short lifespan of the introduced cells, a large number of the corrected T cells have persisted in the first patient for over 6 months after cell infusions were stopped.

The second gene therapy trial involving correction of single-gene hereditary defects is an attempt to restore LDL receptor function in patients with defective or absent LDL receptors by gene transfer into hepatocytes. Unlike the case of ADA deficiency, a suitable animal model for this disorder exists, the Watanabe hyperlipidaemic rabbit. Recent experiments showed that transfer of the rabbit LDL receptor into rabbit hepatocytes in short-term culture, followed by reintroduction of these cells into the liver, resulted in a 30% average reduction in cholesterol levels in the circulation for up to 4 months⁶. These results are very encouraging for the clinical use of these techniques in humans.

The second class of gene therapy experiments involves the transfer of genes to stimulate immunity against or otherwise cause the destruction of tumour cells. The first experiment of this type involved transfer of a tumour necrosis factor (TNF) gene to tumour-infiltrating lymphocytes, in the hope that localized secretion of TNF by these cells, which home into tumour deposits, would enhance immune destruction of tumour cells. Part of the rationale for this experiment is the observation that systemic levels of TNF that cause toxicity in humans are well below the levels that are effective against tumour cells in mouse models, and therefore if the TNF could be synthesized at high concentrations locally it might be effective while avoiding systemic toxicity.

Experiments have been initiated to transfer genes encoding TNF or interleukin-2 (IL-2) into tumour cells in patients in the hope that secretion of these cytokines will stimulate a tumour-specific immune response that would either result in tumour destruction at other sites or allow the collection of more effective TIL from lymph nodes near the site of the injected tumour cells. These experiments are an outgrowth of results from many groups that have shown anti-tumour responses in mice to tumour cells secreting cytokines, including interleukins 2 and 4, γ -interferon, granulocyte colony-stimulating factor, and tumour necrosis factor α (refs 7–14). These studies generally show that injection of cytokine-secreting cells along with parental tumour cells at the same site results in killing of both the modified and unmodified tumour cells, and that injection of cytokine-secreting cells can generate a lasting immunity against the parental unmodified tumour cells. But only one group has shown that cytokine-secreting tumour cells can be effective in stimulating destruction of previously injected parental tumour cells¹⁴, a property that will be important for treatment of most human patients. Even so, this type of therapy might be useful for patients who are apparently disease-free as a result of other antitumour therapy, who would be treated with cytokine-secreting tumour cells in an attempt to prevent future disease recurrence. Although these studies are encouraging, it remains to be seen how general this effect is with regard to human cancer, especially given the variable results seen in mice.

The first attempt (Table 1) to transfer genes directly into somatic cells of a patient without removing the cells to perform the gene transfer procedure involves the transfer of a human class I major histocompatibility antigen (HLA-B7) into tumour cells of patients who do not express this antigen in the hope of stimulating immunity against the tumour cells. The procedure involves direct injection of liposome/plasmid DNA complexes into tumour masses. If such direct injection techniques are effective, they would provide a much simpler and less costly alternative to current *ex vivo* gene transfer methods.

Another gene therapy trial (Table 1) involves an attempt to treat ovarian cancer by intraperitoneal injection of a human

ovarian carcinoma cell line carrying the herpes simplex virus (HSV) thymidine kinase gene. The cells will be lethally irradiated (10,000 rad) before injection to prevent further growth of the cells. At the same time, the patients will be treated with ganciclovir, which can be metabolized by the thymidine kinase gene to a cellular toxin. Anti-tumour effects are presumed to occur by the effect of this toxin on nearby patient ovarian cancer cells, and by stimulation of immunity against tumour cells. Only limited animal data support these concepts, and the mechanism of tumour-cell killing and relevance to treatment of ovarian cancer in humans is far from clear.

Gene transfer vectors

Retroviral vectors. So far all but one of the approved gene transfer trials in humans rely on retroviral vectors for gene transduction, thus it is important to understand the advantages and limitations of this system. Retroviral vectors in this context are retroviruses from which all viral genes have been removed or altered so that no viral proteins are made in cells infected with the vector. Viral replication functions are provided by the use of retrovirus 'packaging' cells that produce all of the viral proteins but that do not produce infectious virus. Introduction of the DNA form of a retroviral vector into packaging cells results in production of virions that carry vector RNA and can infect target cells, but no further virus spread occurs after infection. To distinguish this process from a normal virus infection where the virus continues to replicate and spread, the term transduction rather than infection is often used.

The major advantages of retroviral vectors for gene therapy are the high efficiency of gene transfer into replicating cells, the precise integration of the transferred genes into cellular DNA, and the lack of further spread of the sequences after gene transduction. The ability to transfer genes efficiently and stably to target cells, especially primary somatic cells, is not shared by other gene transfer techniques and is the major attraction of retroviral vectors for use in gene therapy. Major disadvantages include the apparent inability of retroviral vectors to infect nondividing cells¹⁵, and an inherent inability to characterize completely the retroviral vector preparations used for gene transduction because retroviral vectors cannot be made synthetically but must be produced by cultured cells. Unlike proteins or other simple compounds, retrovirus vectors are complex mixtures of proteins and nucleic acids and cannot be purified to homogeneity after production. This disadvantage means that vector-producing cell lines must undergo extensive testing for possible adventitious microorganisms, including replication-competent retroviruses. Other contaminants such as cellular RNAs that are packaged into retroviral vectors cannot be removed. Some of these RNAs can be reverse transcribed and integrated in cells transduced with retroviral vectors¹⁶, and only experience in animal models and in humans will determine their possible effects.

Two other potential problems with retroviral vectors warrant discussion, those of insertional mutagenesis and potential helper virus production. Problems with insertional mutagenesis, such as activation of cellular oncogenes, are shared with any gene transfer technique that results in integration of new sequences into the cellular genome, with the possible exception of adeno-associated virus, which preferentially integrates into one region in chromosome 19 (refs 17, 18). Although there are many examples of retroviral activation of cellular oncogenes in mice, these events occur in the context of a spreading infection by replication-competent virus. Whether such events can occur at appreciable rates after infection by replication-defective retroviral vectors remains to be seen.

The potential for production of replication-competent (helper) virus during the production of retroviral vectors remains a concern, although for practical purposes this problem has been solved. In the human trials so far, none of the production batches of retroviral vectors and none of the human patients have tested

positive for helper virus (P. Tolstoshev and R. Overell, personal communications). The potential for helper virus production depends on viral sequences both in the retroviral vector and in the packaging cells used for vector production (reviewed in ref. 19). The packaging cells contain all sequences necessary for viral protein synthesis and the retroviral vector contains the sequences necessary in *cis* for virus transmission, thus recombination between these sequences has the potential to generate helper virus. Particularly important is the avoidance of homologous overlap between the separate sequences. So far, all FDA-approved vectors have been made by using PA317 amphotropic retrovirus packaging cells²⁰. PA317 cells were made by transfection of mouse cells with DNA containing a contiguous sequence encoding all of the retroviral proteins. It is clear that PA317 cells can produce helper virus after introduction of a retroviral vector that has extensive homologous overlap at both ends of this sequence^{21,22}. But use of vectors having little or no overlap with viral sequences in the PA317 cells eliminates helper virus production even by stringent assays that allow for amplification of such events^{21,22}. Other packaging cell designs involve separation of different retroviral coding regions onto different plasmids, which should reduce the possibility of helper virus production by recombination, and vectors produced by such packaging cell lines are currently undergoing review by the Food and Drug Administration.

Recently, three out of eight rhesus monkeys transplanted with bone marrow infected with a retroviral vector preparation that was heavily contaminated with replication competent retrovirus have developed thymic lymphoma after about 6 months²³. The vector used was produced by co-cultivation of packaging cell lines to produce a high vector titre, but helper virus was also generated by this procedure³⁰. Because high vector titre improves the infection rate in bone marrow cells, the helper-contaminated vector was used in an attempt to show bone marrow stem cell infection in monkeys. Disease induction in these monkeys contrasts with previous experiments which did not detect disease in cynomolgous macaque and rhesus monkeys after exposure to amphotropic helper virus^{24,45}, and which were used to support the safety of proposed human gene therapy experiments. The occurrence of disease after exposure of some monkeys to replication-competent retrovirus reinforces the need for stringent tests for the absence of helper virus in vector preparations for use in humans.

The other gene transfer method that has been approved for use in humans is the transfer of plasmid DNA in liposomes directly to tumour cells *in situ*. Plasmid DNA should be easy to certify for use in humans because, unlike retroviral vectors, it can be purified to homogeneity. In this case, stable integration of the DNA into transduced tumour cells is not required for therapy as transient expression may suffice.

Adenovirus vectors. Recent work with gene transfer vectors based on adenovirus suggests that this vector system will soon be an important tool in gene therapy. Major advantages of adenovirus vectors are their potential to carry large segments of DNA (36-kilobase-pair genome), very high titre (10^{11} ml⁻¹), ability to infect nonreplicating cells, and suitability for infecting tissues *in situ*, especially the lung. Major disadvantages are the inclusion of many adenovirus genes in current vectors that may stimulate immunity or have other adverse effects, and potential instability of gene expression because the vector does not integrate into chromosomal DNA. The most striking use of this vector so far is to deliver a human cystic fibrosis transmembrane conductance regulator (CFTR) gene by intratracheal instillation to airway epithelium in cotton rats²⁵. Expression of the transferred CFTR gene was detected for 6 weeks after the procedure by northern analysis, and human CFTR protein was detected by using an anti-human CFTR antibody at 11 to 14 days after infection. These encouraging results suggest that human trials may soon begin. The rapid progress in cloning the gene involved in cystic fibrosis, understanding the disease biology, and the development

of treatments for this disease is truly impressive.

Physical DNA transfer methods. In addition to liposome-mediated DNA transfer, which will be used to transfer antigen-encoding genes into tumours in humans, several other methods for the direct transfer of plasmid DNA into cells show promise in human gene therapy. Both involve targeting the DNA to receptors on cells by complexing the plasmid DNA to proteins. In one method, polylysine is conjugated to an asialoglycoprotein which is then used to form a complex with plasmid DNA. After injection into animals, the complex is specifically targeted to unique receptors on hepatocytes that internalize galactose-terminal (asialo) glycoproteins. Gene expression is transient after transfer, but inclusion of a two-thirds partial hepatectomy step shortly after infusion of complexed DNA leads to hepatocyte replication and prolonged expression of the introduced gene. Using the latter technique, transfer of a gene encoding human serum albumin into Nagase analbuminaemic rats resulted in circulating human albumin levels of about 30 µg ml⁻¹ for over a month²⁶. More recent results show that the expression persists at this level for four months but then declines to undetectable levels by six months (G. Wu, personal communication). Although this level of albumin is 1,000 times lower than normal circulating levels of albumin, for other blood proteins this would be a quite respectable level of protein production. For example, normal levels of clotting factor IX are only 5 µg ml⁻¹. But to extrapolate these results to other serum proteins one must consider the relatively long serum half-life of albumin of about 2 weeks. Factor IX, for example, has a half-life of only 19 h, thus rate of production of factor IX must be about 20-fold higher than albumin to achieve the same plasma concentration.

Another method for physical transfer of DNA to cells involves targeting of the DNA to the transferrin receptor on cells by complexing the DNA with transferrin²⁷. More recently, the method has been dramatically improved by the use of adenovirus to facilitate exit of DNA from endosomes before being targeted to lysosomes for destruction²⁸. This technique allows transient high-level expression of genes in many cell types, and may be useful in providing short-term gene therapy. Refinement of this technique may involve purification of specific adenovirus proteins that could be used to facilitate physical entry of DNA in general.

Target tissues

Haematopoietic stem cells. One of the main targets for gene therapy is the haematopoietic system because of well developed procedures for bone marrow transplantation, the many types and wide distribution of haematopoietic cells, and the existence of many diseases that affect haematopoietic cells. The target for gene transfer is the haematopoietic stem cell, or long-term repopulating cell, that is present at low frequency in bone marrow and gives rise to all myeloid and lymphoid cells over prolonged periods. Although many groups have reported transfer and long-term expression of genes in mouse haematopoietic cells⁴⁶, the same has not been achieved in larger outbred animals. Monkeys are perhaps the best animal model for procedures that will be used in humans because of their phylogenetic proximity, but dogs are also a good model and have been instrumental in the development of many bone marrow transplantation procedures.

Experiments in monkeys have documented expression of human ADA and neomycin phosphotransferase at low levels in peripheral blood cells for up to 4 months after bone marrow infection with a retroviral vector carrying ADA and *neo* genes²⁹. Persistence of vector sequences in about 1% of peripheral blood cells for up to 3 months after retrovirus-mediated gene transfer has also been demonstrated using a high-titre vector carrying the *neo* gene³⁰, and more recent experiments from the same group document still longer persistence of vector sequences in additional animals. But these experiments are complicated by

the presence of helper virus in the vector preparations used for infection, which can lead to spread of the vector in the animal and has caused lymphoma in several animals. Thus the feasibility of gene transfer in monkeys must be re-evaluated by using helper-free vectors for gene transfer.

Early gene transfer experiments into canine haematopoietic cells resulted in only transient presence of vector function in haematopoietic cells³¹. Recent improvements in the vector infection protocol have resulted in persistence of vector sequences in both lymphoid and myeloid cells in several dogs for over two years, indicating successful transduction of haematopoietic stem cells³². In this case, the vectors were helper-virus-free and helper virus has not been detected in blood from animals receiving vector-infected cells, nor have the animals developed vector-related disease. The vectors used carry the human ADA and bacterial *neo* genes, and drug resistance due to expression of the *neo* gene has been observed intermittently in up to 10% of haematopoietic colonies (CFU-GM) grown from bone marrow for two years. Although these results are encouraging, the level of *neo* expression is low and expression of the ADA gene was not detected.

Retroviral vectors that are similar to those used in monkeys have also been used to infect haematopoietic cells in cats³³. Vector sequences and drug-resistant haematopoietic cell colonies were observed in some of these animals for over 2 years. Human ADA in cats infected with a vector that carried the ADA gene was not detected. As in the monkey experiments, all vector preparations were contaminated with helper virus, and helper virus was detected in the experimental animals. Two of the four experimental animals developed diabetes mellitus 90 days after receiving infected marrow, most probably due to the presence of helper virus.

In utero transfer of the *neo* gene into haematopoietic cells of fetal sheep has been accomplished³⁴ by removing circulating cells 7 weeks before birth, infecting these cells *in vitro*, and returning the infected cells to the fetus. Persistence of the infected cells after birth was demonstrated by detection of *neo* sequences in blood, bone marrow, spleen and thymus samples from some animals 1 week after birth, detection of neomycin phosphotransferase in bone marrow of one animal 6 weeks after birth, and by the presence of G418-resistant haematopoietic colony-forming cells in marrow for more than 2 years after birth. The vectors used were the same as those used in the above experiments in monkeys and cats and were contaminated with helper virus. The same infection procedure was not successful with marrow cells from adult sheep instead of fetal cells. This difference is most probably due to the high replication rate of fetal haematopoietic cells, which would make them more susceptible to infection by the retroviral vector.

It is clear from the above studies that more needs to be done to improve the efficiency of gene transfer and expression in large animal models to provide a solid rationale for extending these studies to trials of gene transfer into human haematopoietic stem cells. Alternatively, some researchers are attempting to identify protocols in which the ability to infect human stem cells can be tested directly without undue danger to the patient. For example, in experiments designed to mark cancerous cells in bone marrow to study the source of cancerous cells during relapse, it is possible that bone marrow stem cells will be marked as an incidental result of the procedure. The presence of the vector in a variety of haematopoietic lineages for long periods after the procedure would provide evidence of stem cell infection. The difficulty with this approach is that conditions required for efficient gene transfer into haematopoietic stem cells are likely to be very different from those required for efficient transfer into cancerous cells in marrow.

Another attempt to infect human stem cells has been approved as an amendment to the protocol involving ADA gene transfer to lymphocytes from ADA-deficient patients. I have listed this trial as a separate trial in Table 1 because it represents a concep-

tually different experiment from the first trial for several reasons. Most importantly, modified lymphocytes can be grown in cell culture and gene transfer verified before the cells are returned to the patient, whereas no *in vitro* assays exist for human haematopoietic stem cells, let alone for the rate of gene transduction. In this trial, patients will be pretreated for about a week with G-CSF in an attempt to stimulate the circulation of stem cells that are normally present in bone marrow. Circulating mononuclear cells will be obtained by apheresis and enriched for stem cells by selection for the cell surface marker CD34. The enriched cells will be incubated for a few days with virus in a cocktail of growth factors that is expected to stimulate stem cell replication, and will then be reinfused. Monkeys provide a good model for this experiment because several antibodies against human CD34 also react with monkey CD34, and although there are no ADA-deficient monkeys, the overall strategy of the experiment could be tested. However, the experiment was approved in the absence of compelling animal data, presumably because retroviral transduction of CD34⁺ cells should present little additional danger to a patient who is already receiving large numbers of retrovirus-transduced lymphocytes, and there might be some benefit if it works.

Muscle cells. It has been possible to transfer genes directly into skeletal muscle *in vivo* by direct injection of DNA^{35,36}. Injection of a plasmid containing a β -galactosidase gene into mouse skeletal muscle resulted in β -galactosidase expression in individual myotubes in the area of the injection. Injection of a plasmid containing a luciferase gene resulted in expression of luciferase for 2 months at relatively constant levels, suggesting that expression may persist for longer periods. The injected DNA seemed to persist in an unintegrated extrachromosomal state. Attempts to transfer genes to other tissues by using DNA have not been successful, suggesting that the ability to incorporate and express naked DNA is unique to muscle³⁵. Unfortunately, transfer of the luciferase gene into skeletal muscle of monkeys resulted in much lower levels of luciferase expression than that observed in mice³⁷. It will be interesting to see if this technique can be developed to provide long-term gene expression at useful levels, as the technique would provide an exciting alternative to the more complex methods for gene transfer currently in use.

Recent experiments have demonstrated that genetically modified myoblasts can be injected into skeletal muscle, fuse with existing muscle and continue to express the transferred gene^{38,39}. After transfer of genes encoding human growth hormone, growth hormone was detected in the serum of recipient animals at levels from 1 to 16 ng ml⁻¹ for up to 85 days. But the C2C12 myoblast cell line used in these experiments is an immortal mouse myoblast line that can form tumours in recipient animals. Indeed, the level of growth hormone detected in the longest experiments increased with time, suggesting that the implanted cells were continuing to grow³⁹. These results are reminiscent of early studies with immortal mouse fibroblast cell lines (NIH 3T3 cells) that were modified to express human factor IX. After subcutaneous or intraperitoneal injection, the cells continued to grow and circulating levels of human factor IX reached 500 ng ml⁻¹ plasma⁴⁰. Normal fibroblasts behave quite differently and can suppress expression of transferred genes after implantation⁴¹. Thus it will be interesting to see if normal myoblasts continue to express introduced genes at meaningful levels after implantation, and whether these techniques can be applied to large animals.

In addition to skeletal muscle myoblasts, smooth muscle cells present in blood vessels can also be cultured, genetically modified and returned to blood vessels after disruption of the vessel wall⁴². Normal rat smooth muscle cells were transduced with a retroviral vector carrying human ADA and bacterial *neo* genes and were returned to the arterial wall after denudation of endothelial cells by passage of a balloon catheter. Human ADA was expressed in the vessel wall for at least 6 months at

TABLE 2 Diseases caused by single gene defects: current targets for gene therapy

Disease	Defective gene
Immunodeficiency	Adenosine deaminase
	Purine nucleoside phosphorylase
Hypercholesterolaemia	LDL receptor
Haemophilia	Factor IX
	Factor VIII
Gaucher's disease	Glucocerebrosidase
Mucopolysaccharidosis	β -glucuronidase
Emphysema	α 1-antitrypsin
Cystic fibrosis	Cystic fibrosis transmembrane regulator
Phenylketonuria	Phenylalanine hydroxylase
Hyperammonaemia	Ornithine transcarbamylase
Citrullinaemia	Arginosuccinate synthetase
Muscular dystrophy	Dystrophin
Thalassaemia	β -globin
Sickle cell anaemia	β -globin
Leukocyte adhesion deficiency	CD-18

This list indicates the principal current targets for gene therapy efforts and is not meant to be complete. Many of the diseases listed can be caused by defects in more than one gene; the gene defect listed is the defect targeted by current research.

relatively constant levels that were similar to the endogenous rat ADA levels in the vessel.

Disease targets

Research in gene therapy has focused on a variety of diseases (Table 2), most of which involve recessive single-gene disorders that can be corrected by addition of a functioning gene to the appropriate cells. In cases like the clotting factor deficiencies, almost any somatic cell can be a target for gene therapy because the gene product is secreted and is required systemically. In other cases the gene must be delivered to a specific cell type, such as the hepatocyte in the case of the urea cycle enzyme ornithine transcarbamylase.

An exciting application of gene therapy will be the treatment of acquired diseases such as cancer and infectious disease. As discussed above, trials for cancer treatment by cytokine gene transfer have already begun. Recent results indicate that AIDS may also be an early target for gene therapy. The focus of such an approach would be to use gene transfer to protect cells from infection by HIV or to prevent the production of HIV from a cell that is already infected. Of the ways that can be imagined to achieve this end, an approach using 'TAR decoys' looks very promising⁴³. In this approach, copies of the short region of HIV that is responsive to the viral transactivator (*tat*) protein, called the transactivation response (TAR) element, are overproduced by including the TAR element in a transfer RNA transcription unit. Transcription by RNA polymerase III leads to very high

levels of the transcripts, which presumably act to bind the *tat* protein and reduce its ability to activate HIV. Cells expressing the TAR decoys could be infected by HIV but would make only low levels of HIV proteins or virus.

Future directions

I have tried to summarize approaches that have the best potential for application to human gene therapy. Many other gene transfer techniques and target tissues are available, and we are likely to see the development of important new techniques in the future. Several areas of research will be instrumental in further progress.

A better understanding of somatic cell transplantation and the biology of the cells that contribute to reconstitution will be necessary. For example, if human haematopoietic stem cells could be expanded in number in culture without diminution of their ability to repopulate the haematopoietic system, as is possible in mice to some extent⁴⁷, such a development would revolutionize the treatment of disease. Then individual stem cell clones could be genetically modified (possibly by homologous recombination), expanded, and used for transplantation. This would be analogous to current techniques for genetic manipulation of embryonic stem cells that can reconstitute whole animals.

The development of improved gene transfer techniques, especially for *in situ* gene delivery, are clearly important. A major problem with most techniques is the inability to stably transfer genes to non-dividing cells. For example, although the elegant technique for targeting DNA to liver by complexation with asialoglycoproteins can result in efficient gene transfer *in vivo*, expression is short lived without a two-thirds hepatectomy to stimulate cell division and (presumably) genomic integration of the DNA.

The consequences of introduction of cells that in general produce proteins that are foreign to the host must be better understood, and methods to control immune response will need development. For example, production of normal LDL receptors in hepatocytes of patients that were previously receptor negative might generate a strong immune response against this new antigen. Even intracellular proteins may elicit detrimental immune responses.

In addition, we must have a better understanding of the factors that control expression of genes introduced into somatic cells. In my own work, I have observed a dramatic decrease in expression (>1,500-fold) of a retroviral vector in skin fibroblasts after transplantation, but not in culture⁴¹, whereas the same vector promotes long-term protein production in smooth muscle cells in the same animal model⁴². The enthusiastic climate surrounding gene therapy should lead to more concentrated efforts in all of these important areas of research. □

A. Dusty Miller is at the Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104, USA.

1. Hum. Gene Ther. **1**, 93-103 (1990).
2. Hum. Gene Ther. **2**, 251-256 (1991).
3. Rosenberg, S. A. Cancer Res. **51** (suppl.), 5074-5079 (1991).
4. Rosenberg, S. A. et al. New Engl. J. Med. **323**, 570-578 (1990).
5. Lupton, S. D., Brunton, L. L., Kalberg, V. A. & Overell, R. W. Molec. cell. Biol. **11**, 3374-3378 (1991).
6. Chowdhury, J. R. et al. Science **254**, 1802-1805 (1991).
7. Tepper, R. I., Pattengale, P. K. & Leder, P. Cell **57**, 503-512 (1989).
8. Watanabe, Y. et al. Proc. natn. Acad. Sci. U.S.A. **85**, 9456-9460 (1989).
9. Fearon, E. R. et al. Cell **60**, 397-403 (1990).
10. Gansbacher, B. et al. J. exp. Med. **172**, 1217-1224 (1990).
11. Gansbacher, B. et al. Cancer Res. **50**, 7820-7825 (1990).
12. Colombo, M. P. et al. J. exp. Med. **173**, 889-897 (1991).
13. Asher, A. L. et al. J. Immun. **146**, 3227-3234 (1991).
14. Golumbek, P. T. et al. Science **254**, 713-716 (1991).
15. Miller, D. G., Adam, M. A. & Miller, A. D. Molec. cell. Biol. **10**, 4239-4242 (1990).
16. Scaden, D. T., Fuller, B. & Cunningham, J. M. J. Virol. **64**, 424-427 (1990).
17. Kotin, R. M. et al. Proc. natn. Acad. Sci. U.S.A. **87**, 2211-2215 (1990).
18. Samulski, R. J. et al. EMBO J. **10**, 3941-3950 (1991).
19. Miller, A. D. Hum. Gene Ther. **1**, 5-14 (1990).
20. Miller, A. D. & Buttimore, C. Molec. cell. Biol. **6**, 2895-2902 (1986).
21. Miller, A. D. & Rosman, G. J. Biotechniques **7**, 980-990 (1989).
22. Lynch, C. M. & Miller, A. D. J. Virol. **65**, 3887-3890 (1991).
23. Kolberg, R. J. NIH Res. **4**, 43-44 (1992).
24. Cornetta, K. et al. Hum. Gene Ther. **2**, 215-219 (1991).
25. Rosenfeld, M. A. et al. Cell **68**, 143-155 (1992).
26. Wu, G. Y. et al. J. biol. Chem. **266**, 14338-14342 (1991).
27. Zenke, M. et al. Proc. natn. Acad. Sci. U.S.A. **87**, 3655-3659 (1990).
28. Curiel, D. T., Agarwal, S., Wagner, E. & Cotten, M. Proc. natn. Acad. Sci. U.S.A. **88**, 8850-8854 (1991).
29. Kantoff, P. W. et al. J. exp. Med. **166**, 219-234 (1987).
30. Bodine, D. M. et al. Proc. natn. Acad. Sci. U.S.A. **87**, 3738-3742 (1990).
31. Stead, R. B., Kwok, W. W., Storb, R. & Miller, A. D. Blood **71**, 742-747 (1988).
32. Schuening, F. G. et al. Blood **78**, 2568-2576 (1991).
33. Lothrop, C. D. et al. Blood **78**, 237-245 (1991).
34. Kantoff, P. W. et al. Blood **73**, 1066-1073 (1989).
35. Wolff, J. A. et al. Science **247**, 1465-1468 (1990).
36. Wolff, J. A. et al. Biotechniques **11**, 474-485 (1991).
37. Jiao, S. et al. Hum. Gene Ther. **3**, 21-33 (1992).
38. Barr, E. & Leiden, J. M. Science **254**, 1507-1509 (1991).
39. Dhawan, J. et al. Science **254**, 1509-1512 (1991).
40. Palmer, T. D., Thompson, A. R. & Miller, A. D. Blood **73**, 438-445 (1989).
41. Palmer, T. D., Rosman, G. J., Osborne, W. R. A. & Miller, A. D. Proc. natn. Acad. Sci. U.S.A. **88**, 1330-1334 (1991).
42. Lynch, C. M. et al. Proc. natn. Acad. Sci. U.S.A. **89**, 1138-1142 (1992).
43. Sullenger, B. A., Gallardo, H. F., Ungers, G. E. & Gilboa, E. Cell **63**, 601-608 (1990).
44. Hum. Gene Ther. **2**, 331-358 (1991).
45. Cornetta, K. et al. Hum. Gene Ther. **1**, 15-30 (1990).
46. Miller, A. D. Blood **76**, 271-278 (1990).
47. Fraser, C. C., Szilvassy, S. J., Eaves, C. J. & Humphries, R. K. Proc. natn. Acad. Sci. U.S.A. **89**, 1968-1972 (1992).

Glacial-to-interglacial variations in the carbon isotopic composition of atmospheric CO₂

Bruno D. Marino, Michael B. McElroy, Ross J. Salawitch
& W. Geoffrey Spaulding*

Department of Earth and Planetary Sciences, and Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

* Dames and Moore, 4220 South Maryland Parkway, Las Vegas, Nevada 89119, USA

Samples of the C4 shrub *Atriplex confertifolia* recovered from packrat middens in the western United States provide a record of changes in the ¹³C/¹²C ratio (δ¹³C) of atmospheric carbon dioxide. During the last ice age, atmospheric CO₂ was isotopically light relative to interglacial periods, a result attributed to a combination of factors including reduced terrestrial biomass and decreased productivity of the polar ocean.

MEASUREMENTS from ice cores¹⁻⁵ have yielded an impressive record of changes in the abundance of atmospheric CO₂ over the past 160,000 years. The concentration of CO₂ was low during glacial times, ~200 p.p.m.; it rose rapidly at the end of the last two ice ages, reaching ~280 p.p.m. during subsequent interglacial epochs. The concentration seems to have been relatively constant for the past 10,000 years, close to the recent pre-industrial value of ~275 p.p.m.

In attempts to account for the changes in CO₂ associated with the transition from glacial to interglacial time (for a review see ref. 6), two general classes of models have been proposed. One invokes an increase in the efficiency of the marine biological pump⁷, due to increased supply of nutrients to the glacial ocean⁸, changes in the Redfield ratio permitting enhanced uptake of carbon for a given supply of nutrients⁸, or an increase in the efficiency with which nutrients are used by the biota at high latitudes reflected in a lower concentration of preformed nitrogen and phosphorus⁹⁻¹¹. The second argues for an increase in the concentration of CO₃²⁻ in glacial surface water, responding to changes in either the abundance or the distribution of alkalinity^{6,12-14}. Models emphasizing an enhanced biological pump in polar regions require values for δ¹³C of atmospheric CO₂ (δ_a¹³) to be higher during glacial than during interglacial conditions⁶. Alkalinity-based models permit lower values of δ_a¹³ for glacial time reflecting, in part, enhanced supply of isotopically light carbon to the atmosphere and ocean from a depleted terrestrial biosphere⁶. We shall argue that constraints imposed by the combination of data for CO₂ and δ_a¹³ require a hybrid model. The reduction in CO₂ in glacial time is attributed to an increase in CO₃²⁻ concentration in surface waters resulting from changes in alkalinity associated with a drop in the level of the lysocline, combined with the effects of a more efficient marine biological pump at low latitudes induced by enhanced supply of nutrients to the surface. Lower values of δ_a¹³ during glacial time, we suggest, are due to a reduction in the quantity of carbon stored as terrestrial biomass combined with a decrease in the productivity of the polar ocean.

Observations

Theory¹⁵ and observation¹⁶ suggest that C4 plants provide a reliable record of the isotopic composition of carbon in atmospheric CO₂. The isotopic composition of carbon in C4 plants (δ_p¹³) is related to δ_a¹³ by a simple transfer function of the form

$$\delta_p^{13} = \delta_a^{13} + \Delta \quad (1)$$

where Δ represents the characteristic difference between δ_a¹³ and δ_p¹³ for a given plant species.

We showed previously that measurements of δ_p¹³ for *Zea mays* could be used to infer accurate values for δ_a¹³, using a value for

Δ based on relations developed by Farquhar¹⁵. Analyses of kernels of *Zea mays* for the period 1948 to 1986, with Δ = 3.27‰, provided results for δ_a¹³ in excellent agreement with data from ice cores and modern air¹⁶. In the absence of relevant physiological data for *Atriplex confertifolia*, we use measurements of δ_p¹³ for modern plants in combination with direct measurements of δ_a¹³ to obtain an empirical estimate for Δ. The value of Δ so derived is applied to analysis of fossil samples of *A. confertifolia*.

Atriplex confertifolia is a spring-blooming desert shrub, distributed widely over both warm and cold deserts of North America¹⁷. Values of δ_p¹³ were obtained from analysis of cellulose nitrate from *A. confertifolia* grown at various locations for the years 1981 (N = 1), 1984 (N = 1) and 1988 (N = 12), where N denotes the number of independent samples. Samples consisted in each case of leaves from five plants growing in close proximity¹⁸. Results were compared with annual mean values of δ_a¹³ obtained as an average of measurements for the Northern Hemisphere reported by Roeloffzen *et al.*¹⁹. Analysis of cellulose nitrate from 14 samples covering all three years indicated a mean value for Δ of 4.42‰ with a standard deviation (s.d.) of ±0.03‰. Plants for 1988, from a diversity of sites ranging from California to Wyoming, yielded a mean value for δ_p¹³ of -12.30 ± 0.11‰. The isotopic composition of the plants was remarkably constant, despite large variations in climate and physiography. The observations provide strong support for the pattern expected from physiological theory: the isotopic composition of C4 plants in general and *A. confertifolia* in particular should be insensitive to conditions in the physical environment; they should accurately record, with a relatively constant offset, the prevailing average value of δ_a¹³.

This conclusion is further supported by analyses of *A. confertifolia* samples dating back to 1843. Values of δ_a¹³ inferred in this case using Δ = 4.42‰ agreed with the trend implied by ice-core measurements⁵ and with independent data obtained from analyses of historic samples of *Zea mays*.

The result for Δ inferred for *Atriplex* is consistent with values of 0.36 and 0.44 for the key variables, φ and p_i/p_a, of the physiological theory relating δ_p¹³ to δ_a¹³ for C4 plants: φ defines the bundle sheath leakage factor, and p_i/p_a is the ratio of internal to external partial pressures of CO₂. The values for φ and p_i/p_a inferred here are consistent with results derived from studies of dry matter for several C4 dicotyledonous plants (dicots) investigated by Farquhar¹⁵ and Hattersley²⁰. The results for φ differs significantly, however, from the value of 0.2 obtained by Henderson *et al.*²¹ from measurements of gas exchange for a diverse group of C4 monocots and dicots. The discrepancy could reflect effects of fractionation after photosynthesis. In that case, we believe that dry-matter determination should be more

appropriate for the present application. Further work is required to resolve this issue.

Samples analysed in this study, dating back to ~30,000 yr BP, were recovered from packrat middens in the southwestern United States. They consisted of well preserved leaf material in macrofossil assemblages²². Typically, samples were recovered from well indurated midden sections lacking evidence of post-depositional alteration or contamination. The discrete macrofossil assemblages from which *Atriplex* was derived represent species expected to be present during a given time period. For example, midden samples for the Middle Wisconsin through early Holocene were dominated by extralocal species, such as juniper (*Juniperus osteosperma*), and lacked most of the modern desert thermophiles, such as creosote bush (*Larrea divaricata*), that are characteristic of modern middens. Although radiocarbon dates were not obtained directly for the samples analysed here, the use of associated dates seems justified^{22,23}.

Decomposition by microbial action or chemical degradation, either during the active habitation of the midden by packrats

TABLE 1 Samples and results for $\delta^{13}\text{C}$ of fossil *Atriplex confertifolia* cellulose nitrate

Date*	Site†	$\delta_p^{13}\text{C}(\text{‰})\ddagger$
~200	Skeleton Hills 3	-10.744
~200	Horse Thief Hills 2	-10.816
~200	Specter Range 2	-10.671
~200	Eureka View 2 A(1)	-10.682
1,580 ± 140	Eureka View 2B	-10.501
2,635 ± 140	Eureka View 1	-10.867
5,595 ± 210	Eureka View 4(1)1	-10.958
6,795 ± 190	Eureka View 4(1)2	-10.992
10,690 ± 280	Horse Thief Hills 1	-10.725
12,770 ± 150	Point of Rocks 3(7)	-11.319
14,720 ± 530	Eureka View 5A	-11.138
14,810 ± 400	Point of Rocks 3(5)	-11.831
17,000 ± 270	Point of Rocks 3(2)1	-11.961
20,200 ± 1,800	Specter Range 3 A(1)	-11.512
25,900 ± 260	Specter Range 2 (6)	-11.519
28,460 ± 2,200	Specter Range 3 B	-11.427

* Radiocarbon years BP as described in ref. 22.

† Site locations²² (latitude, longitude, m above sea level): Eureka View, 37° 20' N, 117° 47' W, 1,460 m; Horse Thief Hills, 37° 21' N, 117° 48' W, 1,635 m; Skeleton Hills, 36° 38' N, 116° 17' W, 940 m; Specter Range, 36° 40' N, 116° 13' W, 1,190 m; Point of Rocks, 36° 34' N, 116° 05' W, 910 m.

‡ Measured $\delta^{13}\text{C}$ of *Atriplex*, with respect to PDB, where $\delta = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}} - 1]$. The isotopic analyses reported here were done on a VG PRISM mass spectrometer at Hoffman Laboratory, Harvard University. An auto-cryogenic cold finger was used for small samples. Measurements of $\delta^{13}\text{C}$ were made from analyses of CO_2 obtained by combustion of cellulose nitrate (prepared as in ref. 50) to ensure a contaminant free substrate: seven replicate analyses of *Atriplex* cellulose nitrate yielded a mean value for the differences in $\delta^{13}\text{C}$ of $0.06 \pm 0.05\text{‰}$. An internal standard (pure α -cellulose) included in 11 separate sets of nitrations yielded a mean and s.d. of $-23.82 \pm 0.02\text{‰}$. Over the period of the study we analysed 15 replicate CO_2 tank gas standards, including samples of varying sizes in addition to the cryogenically trapped samples, to check for drift: the analyses yielded a mean and s.d. of $-45.54 \pm 0.04\text{‰}$. Analyses of NBS 22 oil ($\delta^{13}\text{C} = -29.61\text{‰}$) and NBS 19 marble ($\delta^{13}\text{C} = +1.93\text{‰}$) samples during the period of the study yielded means and s.d.s of $-29.66 \pm 0.09\text{‰}$ ($N=5$) and $+1.96 \pm 0.02\text{‰}$ ($N=4$), respectively ($\delta^{13}\text{C}$ values in brackets are NBS reference values⁵¹). On the basis of these data, we estimate the analytical uncertainty of our measurements as better than 0.10‰. Taking this in conjunction with the data on natural variability described in the text, we estimate an overall external precision for the present results of ~0.11‰.

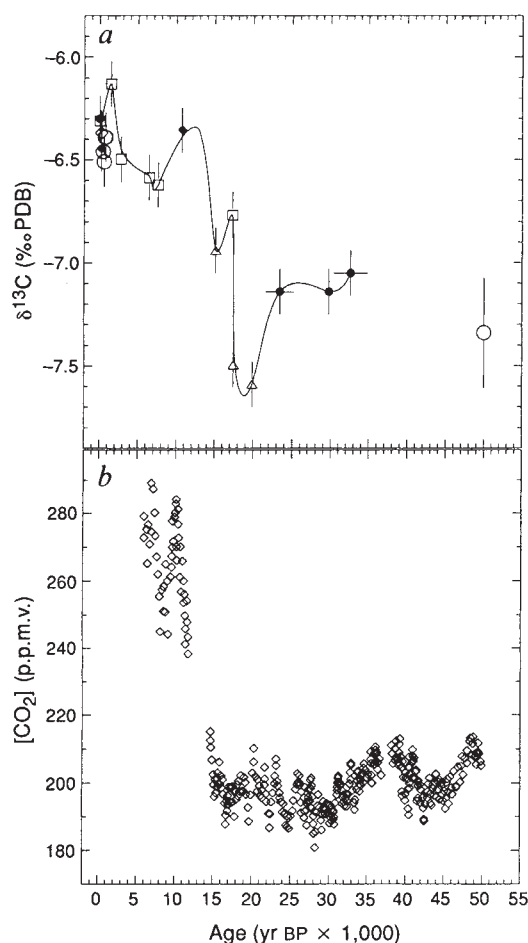


FIG. 1 a, Values of $\delta_a^{13}\text{C}$ as a function of time (calendar years BP) inferred from measurements of $\delta^{13}\text{C}$ in cellulose nitrate of *Atriplex confertifolia*. Each point represents a composite leaf sample from a discrete, dated stratigraphic unit as described in Spaulding²² for the following sites: Skeleton Hills ($\diamond$); Horse Thief Hills ($\blacklozenge$); Eureka View ($\square$); Point of Rocks ($\triangle$); Specter Range ($\bullet$). The radiocarbon ages and uncertainties are given in Table 1. Calibrated ages were determined as described in ref. 52 for samples younger than 8,000 ^{14}C yr BP; the relation between ^{14}C and Th/U ages described in Fairbanks²⁵ was used to calibrate older samples. Larger open circles ($\circ$) represent the mean and standard deviation of $\delta_a^{13}\text{C}$ obtained from Antarctic ice against the mean age of the gas, which has an estimated uncertainty of 33% (ref. 26). The values have been corrected by +0.24‰ for interference by N_2O (ref. 28). b, CO_2 concentration from the Byrd ice core, Antarctica ($\diamond$) as a function of time. Concentration values and timescale are as reported in ref. 4.

or associated with the subsequent postdepositional history, could have affected the values of $\delta_p^{13}\text{C}$. Given the well preserved nature of the samples and results of laboratory studies of cellulose degradation²⁴, we consider it unlikely that values of $\delta_p^{13}\text{C}$ could have been altered by diagenesis.

Values for $\delta_a^{13}\text{C}$ are presented in Fig. 1a as a function of time before present (yr BP). Radiocarbon ages were corrected for secular changes in production of ^{14}C using the relation between ^{14}C and Th/U ages reported by Fairbanks²⁵. Information on individual isotopic determinations, radiocarbon ages and provenance is summarized in Table 1. Also shown are results for $\delta_a^{13}\text{C}$ obtained for CO_2 extracted from ice at the South Pole²⁶. The CO_2 record⁴ is presented in Fig. 1b to facilitate comparison of temporal trends for CO_2 and $\delta_a^{13}\text{C}$. Results for $\delta_a^{13}\text{C}$ during the Holocene, spanning an interval from the pre-industrial era to ~11,000 ^{14}C yr BP, indicate a range of values from -6.13 to -6.62‰, corresponding to a mean and s.d. of -6.40 and 0.15‰, respectively. Two samples within the transition period average $-6.86 \pm 0.12\text{‰}$. Data for the three oldest samples (20,200 to 28,460 ^{14}C yr BP), expected to represent full glacial conditions, indicate an average of $-7.11 \pm 0.05\text{‰}$. The rise in $\delta_a^{13}\text{C}$ marking the early Holocene seems to be preceded by an interval of low $\delta_a^{13}\text{C}$, with values of -7.46 and -7.59‰ indicated for a pair of samples dated at 14,800 and 17,000 ^{14}C yr BP, respectively. If the average Holocene value for $\delta_a^{13}\text{C}$ is taken as -6.43‰, the

difference in δ_a^{13} between glacial and interglacial conditions is estimated as 0.7 to -1% .

Values of δ_a^{13} obtained here for the Holocene are in good agreement with results reported²⁶ for air trapped in ice at the South Pole, where the average value was $-6.44 \pm 0.08\%$ (with a correction of $+0.24\%$ for interference by N_2O) for air with average ages ranging from 430 to 770 yr (Fig. 1a). Measurements of δ_a^{13} for glacial air extracted from the Byrd core are also consistent with the present results: with an estimated age of 50,000 yr, these data indicate a range of values from -7.13 to -7.64% (applying a similar correction for N_2O). Friedli *et al.*²⁶ caution "that the quality of the Byrd ice samples was not favourable and that therefore the results are less reliable than those of the recently obtained South Pole samples". Their results were not corrected for gravitational fractionation²⁷ or for fractionation associated with the extraction procedure. These effects, however, are of opposite sign and partially offset each other.

The glacial-to-interglacial change in δ_a^{13} implied by the present results is less than that based on analysis of CO_2 extracted from ice cores from Byrd Station, Antarctica²⁸. Those results suggest that glacial values of δ_a^{13} were lower than interglacial values by $0.3 \pm 0.2\%$. The nature of the discrepancy is unclear: it could reflect environmental influences on Δ and consequent uncertainty in the values of δ_a^{13} derived from the fossil plants; alternatively, it could reflect analytical difficulties associated with measurements of the isotopic composition of CO_2 in ice. As discussed above, the relatively small range of values for δ_p^{13} observed for contemporary samples of *Atriplex* from diverse environments supports the assumption of a constant value for

Δ . In the latter case, the relatively large range of values implied by the ice-core measurements for glacial δ_a^{13} (-6.44 to -7.10%), and the unexplained analytical problem which requires a correction to values derived for δ_a^{13} (0.21% , assumed constant), suggest that there may be ambiguities in the interpretation of the ice data comparable at least to those for plants.

Errors in derivation of δ_a^{13} from plant proxies could arise if δ_p^{13} reflected physiological responses induced by climate change rather than variations in δ_a^{13} . The value for ϕ (0.36) implied by the present analysis for *Atriplex* suggests minimal changes in δ_p^{13} in response to variations in p_i/p_a ; for the limiting case $\phi = 0.37$, δ_p^{13} is independent of p_i/p_a (refs 15, 16). Lower values of p_i/p_a could arise as a consequence of vapour pressure deficits resulting from increased aridity, leading to stomatal closure; this would result in lighter values for δ_p^{13} and thus heavier values for δ_a^{13} for ϕ less than 0.37. In contrast, reduced concentrations of CO_2 or impaired photosynthetic function could result in an increase in stomatal aperture, leading to slightly higher values for p_i/p_a and δ_p^{13} , and thus lighter values for δ_a^{13} for $\phi < 0.37$.

Further work is needed to characterize the plant approach fully and to resolve questions relating to the interpretation of the ice data. The two approaches agree in the main conclusion that carbon in atmospheric CO_2 was isotopically lighter during glacial than during interglacial time.

Model studies

Models addressing differences between the glacial and interglacial environments must account for several constraints imposed by data from deep-sea sediments and polar ice (see,

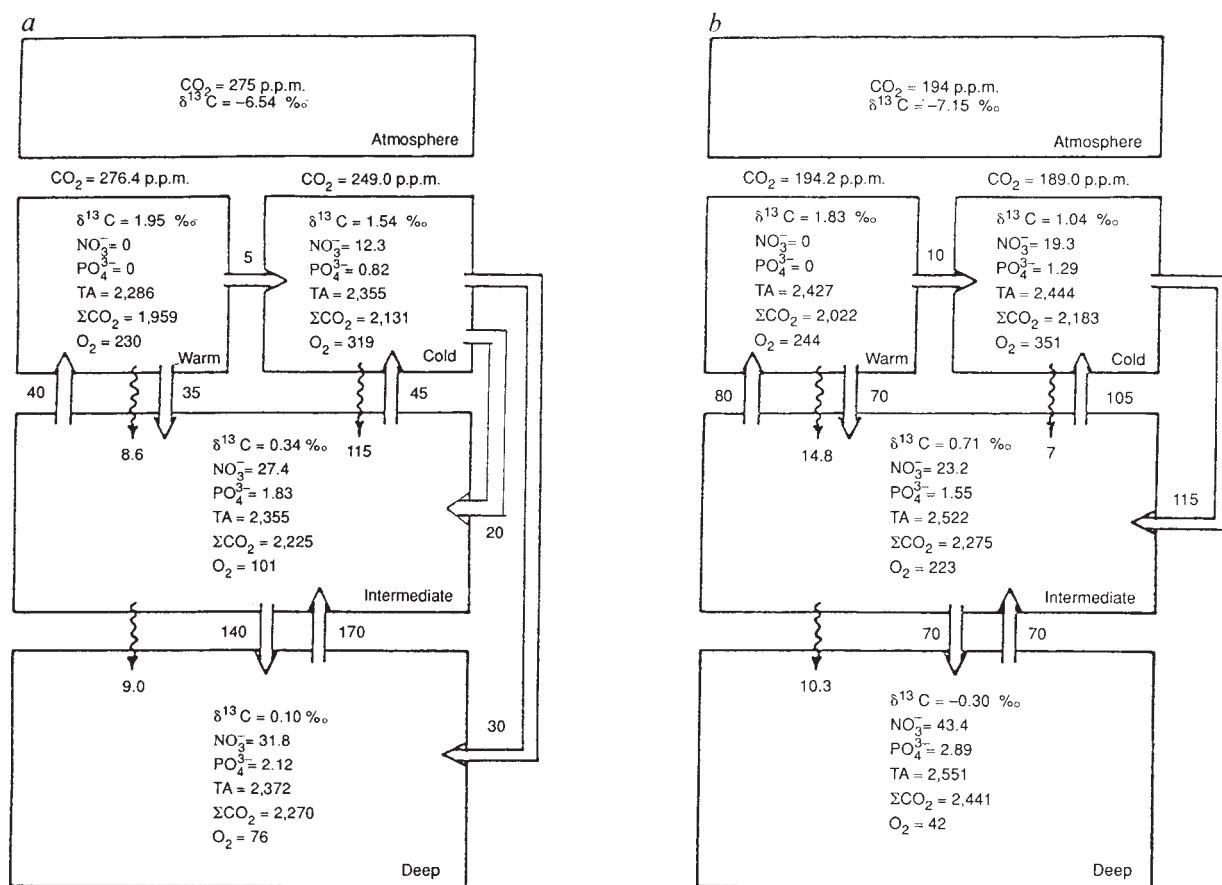


FIG. 2 a, Five-box kinetic model for CO_2 and ^{13}C . Advective fluxes of water (indicated by double arrows) are given in Sverdrups (10^6 m³ s⁻¹). Calculated concentrations and fluxes of biogenic carbon (indicated by wavy arrows) are given in $\mu\text{mol kg}^{-1}$ and $\text{g C m}^{-2} \text{yr}^{-1}$, respectively. The partial pressure of CO_2 above the warm and cold surface water before atmospheric transfer

are indicated. Results correspond to standard interglacial conditions (model A). Further details are given in Table 2 and the text. b, As a, except for glacial model G. Advective fluxes correspond to the shallower, more-vigorous thermohaline circulations adopted for glacial models. Further details are given in Table 2 and the text.

TABLE 2 Summary of model results

Model	Model inputs (average ocean)		Results						
	[TA] $\mu\text{mol kg}^{-1}$	$[\Sigma \text{CO}_2]$ $\mu\text{mol kg}^{-1}$	$[\text{CO}_2]$ atmos (p.p.m.)	$\delta^{13}\text{C}$ atmos (‰)	$\delta^{13}\text{C}$ WSW (‰)	$\delta^{13}\text{C}$ CSW (‰)	$\delta^{13}\text{C}$ IW (‰)	$\delta^{13}\text{C}$ DW (‰)	CaCO_3 preservation rate ($10^{13} \text{ mol yr}^{-1}$)
Interglacial									
A	2,365	2,250	275	-6.54	1.95	1.54	0.34	0.10	2.2
Glacial									
B	2,480	2,372	281	-7.17	1.56	1.19	-0.06	-0.30	2.2
C	2,479	2,348	250	-6.71	2.00	1.53	0.94	-0.30	2.2
D	2,530	2,374	236	-6.77	1.92	1.46	0.90	-0.30	3.7
E	2,540	2,377	197	-6.50	2.18	1.65	1.02	-0.30	4.4
F	2,543	2,381	220	-6.82	1.87	1.41	0.87	-0.30	1.7
G	2,540	2,384	194	-7.15	1.83	1.04	0.71	-0.30	2.4
H	2,540	2,384	193	-7.49	1.69	1.13	0.72	-0.30	2.4
I	2,540	2,384	193	-7.44	1.72	1.11	0.72	-0.30	2.4
J	2,540	2,384	195	-6.92	1.92	0.98	0.71	-0.30	2.4

Abbreviations: TA, titration alkalinity; ΣCO_2 , total dissolved inorganic carbon; atmos, atmosphere; WSW, warm surface water; CSW, cold surface water; IW, intermediate water; DW, deep water. Standard conditions for the interglacial ocean correspond to temperatures of 19 and 2 °C, respectively, for WSW and CSW; salinity = 34.7‰; average oceanic $[\text{NO}_3^-] = 29.1 \mu\text{mol kg}^{-1}$; and $\text{CaCO}_3/\text{C}_{\text{organic}} = 0.25$ and 0.03, respectively, for WSW and CSW. Standard conditions for the glacial ocean correspond to temperatures of 17 and -1 °C, respectively for WSW and CSW; salinity = 35.7‰; average oceanic $[\text{NO}_3^-] = 30.6 \mu\text{mol kg}^{-1}$; and $\text{CaCO}_3/\text{C}_{\text{organic}} = 0.25$ and 0.0, respectively, for WSW and CSW. Film thicknesses for CO_2 and O_2 invasions in warm and cold surface waters are 25 and 20 μm , respectively. The fraction of incoming nutrients removed by photosynthetic organisms in CSW is 0.50, and the fractional area of CSW is 0.04, unless indicated otherwise. For each model, average oceanic $[\Sigma \text{CO}_2]$ is adjusted by the addition or subtraction of a quantity of organic carbon ($\delta^{13}\text{C} = -26\text{‰}$) so that for DW, $\delta^{13}\text{C} = +0.10$ and -0.30‰ for interglacial and glacial conditions, respectively, consistent with the change in $\delta^{13}\text{C}$ of DW implied by observation^{29,30,32}. The amount of carbon transferred from organic to inorganic reservoirs, defined as ΔC , is given in gigatonnes carbon (1 Gt C = 10^9 tonnes) for individual models below. The last column summarizes rates at which CaCO_3 is incorporated in sediment. In steady state, preservation of CaCO_3 in sediments must be balanced by input from continents supplied mainly by rivers.

A: Standard interglacial conditions. The depth of the lysocline is 3,256 m. See Fig. 2a for further details on results.

B: Standard glacial conditions and the addition of 3.6×10^{16} mol of organic C to the ocean, sufficient to lower $\delta^{13}\text{C}$ of DW by 0.4‰. Average oceanic [TA] and $[\Sigma \text{CO}_2]$ have been adjusted so that the depth of the lysocline (3,304 m) yields a preservation rate of CaCO_3 equal to the interglacial value. $\Delta\text{C} = 566$ Gt C.

C: Same as B, except that the oceanic circulation depicted in Fig. 3b has been used. The depth of the lysocline is 3,087 m, yielding a preservation rate of CaCO_3 equal to the interglacial model. $\Delta\text{C} = 186$ Gt C.

D: Same as C, except that average oceanic [TA] and $[\Sigma \text{CO}_2]$ have been adjusted so that the lysocline is 700 m deeper than the interglacial level³⁴. $\Delta\text{C} = 194$ Gt C.

E: Same as D, except that average oceanic $[\text{NO}_3^-] = 36.7 \mu\text{mol kg}^{-1}$. $\Delta\text{C} = 162$ Gt C.

F: Same as D, except that $\text{CaCO}_3/\text{C}_{\text{organic}} = 0.125$ for WSW. $\Delta\text{C} = 202$ Gt C.

G: Same as D, except: average oceanic $[\text{NO}_3^-] = 36.7 \mu\text{mol kg}^{-1}$; $\text{CaCO}_3/\text{C}_{\text{organic}} = 0.125$ for WSW; temperature of WSW = 16.5 °C; salinity of WSW and CSW is 34.6‰; the fraction of incoming nutrients removed by biology in CSW is 0.06; and the fractional area of CSW is 0.08. See Fig. 2b for further details on model results. $\Delta\text{C} = 275$ Gt C for models G-J.

H: Same as G, except that the gas transfer rate between CSW and atmosphere has been doubled (film thickness is 10 μm).

I: Same as G, except that the fractional area of CSW is 0.14.

J: Same as G, except that the fractional area of CSW is 0.04.

for example, ref. 6). We must allow for a reduction of $\sim 0.35\%$ in the value of $\delta^{13}\text{C}$ for the deep glacial ocean with respect to interglacial conditions^{29,30}. Measurements of planktonic foraminifera in the Antarctic Ocean³¹ imply a reduction of 0.7 to 1‰ in $\delta^{13}\text{C}$ for the ice-age surface ocean at high latitudes. Values of $\delta^{13}\text{C}$ for the surface ocean at low latitudes show little variation from glacial to interglacial time³². There is evidence from cadmium and ^{13}C in foraminifera³³ for a change in ocean circulation during glacial time favouring formation of intermediate waters at the expense of deep. There are indications of a variation in the depth of the lysocline (defined as the depth at which ocean water becomes undersaturated in CO_3^{2-} with respect to calcite) under glacial conditions, a considerable deepening in the Pacific³⁴ being accompanied by modest shoaling in the Atlantic³⁵. We assume here a global average drop of 700 m, attaching greater significance to the data for the Pacific. The objective for models is that, while accommodating these constraints, they should account for the changes observed in the concentration and isotopic composition of CO_2 .

We used a box model for the atmosphere-ocean system, as illustrated in Fig. 2a and b. The computational approach is similar to that described by Knox and McElroy^{9,36,37}, accounting for effects of dissolved CO_2 on the isotopic fractionation of phytoplankton carbon³⁸. Results for a series of models are summarized in Table 2. Model A (summarized in Fig. 2a), with CO_2 equal to 275 p.p.m. and $\delta_a^{13}\text{C}$ equal to -6.54‰ , is presented

as a reference for interglacial conditions. The glacial environment is addressed in the subsequent models, which assume in all cases a reduction in ocean volume appropriate for a drop in sea level of 100 m with an associated increase in average ocean salinity, nutrients, dissolved inorganic carbon and alkalinity. The glacial models allow for transfer of carbon (with $\delta^{13}\text{C}$ equal to -26‰) from the biosphere-soil system to the ocean sufficient to bring the value of $\delta^{13}\text{C}$ for deep water to -0.3‰ , accounting for a glacial-to-interglacial difference of 0.4‰.

Model B explores the effects of changes in temperature and sea level on CO_2 and $\delta^{13}\text{C}$. The temperatures of warm and cold surface waters were set equal to 17 and -1 °C, respectively. Model B implies an increase in CO_2 to 281 p.p.m. with a reduction in $\delta_a^{13}\text{C}$ to -7.2‰ . The increase in CO_2 results from the higher abundance of inorganic carbon and the increase in salinity, offset in part by the decrease in temperature. The decrease in $\delta_a^{13}\text{C}$ reflects the input of isotopically light carbon from the biosphere and soil, amplified for the atmosphere by the higher abundance of dissolved CO_2 in the colder surface ocean. The higher concentration of nutrients in model B with respect to model A (resulting from the decrease in ocean volume) leads to a modest increase in biological productivity ($\sim 3\%$), with a drop in the lysocline of ~ 50 m required to maintain a specified constant rate of preservation of CaCO_3 .

The ocean circulation is adjusted in model C to allow better agreement with the observational data for cadmium and ^{13}C

(ref. 33). The physical model adopted here (see Fig. 2b) is characterized by increased upwelling of intermediate water to the surface at low latitudes and by a marked reduction in the rate of ventilation of the deep sea. The shallower, more vigorous, circulation is associated with enhanced biological productivity in response to increased supply of nutrients to the surface ocean with an associated change in the depth distribution of CO_3^{2-} and $\delta^{13}\text{C}$. The abundance of CO_2 is reduced in model C with respect to model B by 31 p.p.m. (to 250 p.p.m.), whereas δ_a^{13} is increased from -7.2 to -6.8% . The value of $\delta^{13}\text{C}$ for the cold surface reservoir (CSW) in model C is similar to that for the interglacial reference model A, in contrast to the depletion observed for the high-latitude glacial surface ocean³¹. Model C is similar to that proposed by Boyle³⁹, except that it allows for an extra source of inorganic carbon. The quantity of carbon transferred from organic to inorganic reservoirs required to effect the postulated changes in $\delta^{13}\text{C}$ for deep water in model C is significantly less than that for model B, 186 as compared with 566 gigatonnes (Gt) carbon, as indicated in Table 2 (see also ref. 40). An additional reduction in CO_2 of nearly 40 p.p.m. is needed to account for the observed glacial value of ~ 200 p.p.m. Further adjustments are required to accommodate the observed glacial reductions in δ_a^{13} and in $\delta^{13}\text{C}$ for CSW.

Model D allows for an increase in the depth of the lysocline of 700 m, consistent with observations for the Pacific³⁴. The level of CO_2 is reduced with respect to model C by 14 p.p.m., to 236 p.p.m., with an increase of 70% in the rate of preservation of CaCO_3 . Minimal changes are predicted for δ_a^{13} and for $\delta^{13}\text{C}$ of CSW.

Further reductions in CO_2 can be accommodated by an increase in the rate of biological productivity at low latitudes, a decrease in the rate of production of CaCO_3 with respect to organic matter in surface waters, or both. Higher productivity for glacial conditions at low latitudes is consistent with data from a variety of sources^{34,41-48}. Boyle and Keigwin³³ observed an increase of $\sim 40\%$ in the ratio of Cd/Ca for benthic foraminifera in the Atlantic during glacial times, interpreted as a proxy for the abundance of phosphorus. There is evidence, however, for a reduction in Cd/Ca in the equatorial Pacific¹³ offsetting the increase in the Atlantic. It is unclear whether the results reported by Boyle¹³ are representative of the entire Pacific, and it is possible that changes in Cd/Ca could arise in response to environmental influences other than changes in phosphorus. We simulated the effects of an increase in productivity for the low-latitude ocean by increasing the average nutrient content of the ocean by 22% (model E); this resulted in a reduction of CO_2 to 197 p.p.m. A 50% decrease in the relative rate for production of CaCO_3 in warm surface waters, with nutrients held fixed at the level in model D, would be responsible for a drop in CO_2 of 16 p.p.m. to 220 p.p.m. (model F). A shift in biology favouring diatoms over organisms with calcareous hard parts would be consistent with evidence for enhanced upwelling during glacial times⁴⁸. Combinations of models E and F provide acceptable results for CO_2 . They fail, however, to account for constraints imposed by the glacial values of δ_a^{13} and $\delta^{13}\text{C}$ for CSW. Results in both cases are unacceptably high.

The decrease in $\delta^{13}\text{C}$ observed in the high-latitude glacial southern ocean implies a considerable reduction in biological productivity for cold surface water. This in itself would produce an increase in the abundance of CO_2 . The increase in CO_2 due to the change in the high-latitude ocean must be offset by a larger reduction induced by changes at low latitudes. This is accommodated in model G (summarized in Fig. 2b) which invokes, simultaneously: enhanced productivity at low latitudes resulting from a higher abundance of nutrients; an increase in the low-latitude surface concentration of CO_3^{2-} caused by a reduction in the relative rate of production of CaCO_3 ; a decrease in the productivity of the high-latitude ocean; a decrease in salinity at the surface offset by higher salinity at depth; a small

additional reduction in the temperature of the warm surface water; and an expansion in the area of the cold surface. The model is consistent with all of the observational constraints on the differences between the glacial and interglacial environments as summarized above: it accounts for a concentration of CO_2 in the glacial atmosphere of 194 p.p.m. and a value of $\delta_a^{13} = -7.15\%$ (see Table 2). Model G implies a reduction in productivity per unit area of CSW by a factor of 8 with respect to interglacial conditions, consistent with the reduction reported by Mortlock *et al.*⁴⁹.

The decrease in δ_a^{13} , with little associated change in CO_2 , near the end of the glacial period implied by the data in Fig. 1 could be attributed to an increase in the rate of gas exchange between the atmosphere and the high-latitude ocean, or, equivalently, to an increase in the area of the high-latitude surface reservoir. These conditions are simulated in models H and I. Model J is similar to model G except that the fractional area of CSW is reduced by a factor of 2. The concentration of CO_2 is unaffected, but δ_a^{13} is increased from -7.15 to -6.92 . Model J would be consistent with the data presented in ref. 28.

Conclusions

Observations of δ_a^{13} combined with data for CO_2 provide important constraints for models of the glacial ocean. Reductions in δ_a^{13} and $\delta^{13}\text{C}$ for CSW during glacial times are ascribed to a reduction in the productivity of CSW combined with a reduction in the value of $\delta^{13}\text{C}$ for the total inorganic carbon content of the atmosphere and ocean, caused by a decrease in the abundance of carbon stored in organic form in the terrestrial biosphere and soil. Lower concentrations of CO_2 during glacial times are attributed to an increased supply of nutrients to the surface of the low-latitude ocean, and to biological and physical influences resulting in an increase in the abundance of CO_3^{2-} in WSW. □

Received 23 January; accepted 5 May 1992.

1. Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **329**, 408-414 (1987).
2. Neftel, A., Oeschger, H., Schwander, J., Stauffer, B. & Zumbund, R. *Nature* **295**, 220-223 (1982).
3. Neftel, A., Moor, E., Oeschger, H. & Stauffer, B. *Nature* **315**, 45-47 (1985).
4. Neftel, A., Oeschger, H., Staffelbach, T. & Stauffer, B. *Nature* **331**, 609-611 (1988).
5. Staffelbach, T., Stauffer, B., Sigg, A. & Oeschger, H. *Tellus* **B43**, 91-96 (1991).
6. Broecker, W. S. & Peng, T.-H. in *NATO Volume on the Global Carbon Cycle* (in the press).
7. Volk, T. & Hoffert, M. I. in *The Carbon Cycle and Atmospheric CO_2 : Natural Variations Archaean to Present*, *Geophys. Monogr.* **32** (eds Sundquist, E. T. & Broecker, W. S.) 99-110 (American Geophysical Union, Washington DC, 1985).
8. Broecker, W. S. *Progr. Oceanogr.* **11**, 151-197 (1982).
9. Knox, F. & McElroy, M. B. *J. geophys. Res.* **89**, 4629-4637 (1984).
10. Sarmiento, J. L. & Toggweiler, R. *Nature* **308**, 621-624 (1984).
11. Siegenthaler, U. & Wenk, T. *Nature* **308**, 624-626 (1984).
12. Berger, W. H. *Naturwissenschaften* **69**, 87-88 (1982).
13. Boyle, E. A. *Nature* **331**, 55-56 (1988).
14. Broecker, W. S. & Peng, T.-H. *Global biogeochem. Cycles* **3**, 215-239 (1989).
15. Farquhar, G. D. *Austr. J. Plant Physiol.* **10**, 205-226 (1983).
16. Marino, B. D. & McElroy, M. B. *Nature* **349**, 127-131 (1991).
17. Osmond, C. B., Björkman, O. & Anderson, D. J. *Physiological Processes in Plant Ecology* (Springer, New York, 1980).
18. Sanderson, S. C., Stutz, H. C. & McArthur, E. D. *Am. J. Bot.* **58**, 54-68 (1990).
19. Roeloffzen, J. C., Mook, W. G. & Keeling, C. D. *IAEA Tech. Rep. SM-313/93* (1990).
20. Hattersley, P. W. *Austr. J. Plant Physiol.* **9**, 139-154 (1982).
21. Henderson, S. A., von Caemmerer, S. & Farquhar, G. D. *Austr. J. Plant Physiol.* (in the press).
22. Spaulding, W. G. *U.S. geol. Surv. Prof. Pap.* 1329 (U.S. Department of the Interior, 1985).
23. Van Devender, T. R. *et al. Nature* **327**, 610-613 (1985).
24. Marino, B. D. & Deniro, M. J. *J. arch. Sci.* **14**, 537-548 (1987).
25. Fairbanks, R. G. *Paleoceanography* **5**, 937-948 (1990).
26. Friedli, H., Moor, E., Oeschger, H., Siegenthaler, U. & Stauffer, B. *Geophys. Res. Lett.* **11**, 1145-1148 (1984).
27. Craig, H., Horibe, Y. & Sowers, T. *Science* **242**, 1675-1678 (1988).
28. Leuenberger, M., Siegenthaler, U. & Langway, C. C. *Nature* **357**, 488-490 (1992).
29. Curry, W. B., Duplessy, J.-C., Labeyrie, L. D. & Shackleton, N. J. *Paleoceanography* **3**, 317-341 (1988).
30. Duplessy, J.-C. *et al. Paleoceanography* **3**, 343-360 (1988).
31. Charles, C. D. & Fairbanks, R. G. in *Geological History of the Polar Oceans: Arctic versus Antarctic* (eds Bleil, U. & Thiede, J.) 519-538 (Kluwer, Dordrecht, 1990).
32. Shackleton, N. J., Hall, M. A., Line, J. & Shuei, C. *Nature* **306**, 319-322 (1983).
33. Boyle, E. A. & Keigwin, L. D. *Earth planet. Sci. Lett.* **76**, 135-150 (1985).
34. Farrell, J. W. & Prell, W. L. *Paleoceanography* **4**, 447-466 (1989).
35. Balsam, W. L. *J. sedim. Petrol.* **53**, 719-731 (1983).
36. Ennever, F. K. & McElroy, M. B. in *The Carbon Cycle and Atmospheric CO_2 : Natural Variations Archaean to Present*, *Geophys. Monogr.* **32** (eds Sundquist, E. T. & Broecker, W. S.) 154-162 (American Geophysical Union, Washington, DC, 1985).
37. Ennever, F. K. thesis, Harvard Univ. (1985).
38. Rau, G., Froelich, P. N., Takahashi, T. & Des Marais, D. J. *Paleoceanography* **6**, 335-347 (1991).
39. Boyle, E. A. *J. geophys. Res.* **93**, 15701-15714 (1988).
40. Keir, R. S. *Global biogeochem. Cycles* **5**, 351-358 (1991).
41. Berger, W. H. *J. foram. Res.* **3**, 187-195 (1973).

42. Johnson, D. A. & Knoll, A. H. *Quat. Res.* **4**, 206–216 (1974).
43. Adelseck, C. G. & Anderson, T. F. *Geology* **6**, 388–391 (1978).
44. Pedersen, T. F. *Geology* **11**, 16–19 (1983).
45. Lyle, M. *et al. Paleoceanography* **3**, 39–59 (1988).
46. Pedersen, T. F., Pickering, M., Vogel, J. S., Southon, N. J. & Nelson, D. E. *Paleoceanography* **3**, 157–168 (1988).
47. Finney, B. P., Lyle, M. W. & Heath, G. R. *Paleoceanography* **3**, 169–189 (1988).
48. Sarthain, M., Winn, K., Duplessy, J.-C. & Fontugne, M. R. *Paleoceanography* **3**, 361–399 (1988).
49. Mortlock, R. A. *et al. Nature* **351**, 220–223 (1991).
50. Green, J. W. in *Methods in Carbohydrate Chemistry III* (ed. Whistler, R. L.) 110–135 (Academic, New York, 1963).
51. Coplen, T. B., Kendall, C. & Hopple, J. *Nature* **302**, 236–238 (1983).
52. Stuiver, M. & Reimer, P. J. *Radiocarbon* **28**, 1022–1030 (1986).

ACKNOWLEDGEMENTS. We thank G. Farquhar and S. Henderson for discussion and for a preprint of their paper with S. von Caemmerer. We acknowledge also comments by M. Bender, R. Keir, J. McCarthy, S. Sanderson, M. Sarthain, U. Siegenthaler and M. Wahlen. This work was supported at Harvard University by the U.S. Department of Energy through the National Institute of Global and Environmental Change and by the NSF.

X-ray analyses of peptide–inhibitor complexes define the structural basis of specificity for human and mouse renins

V. Dhanaraj^{*†}, C. G. Dealwis^{*†}, C. Frazao[†], M. Badasso[†], B. L. Sibanda[†], I. J. Tickle[†], J. B. Cooper[†], H. P. C. Driessen[†], M. Newman[†], C. Aguilar[†], S. P. Wood[†], T. L. Blundell^{†‡}, P. M. Hobart[§], K. F. Geoghegan[§], M. J. Ammirati[§], D. E. Danley[§], B. A. O'Connor[§] & D. J. Hoover^{||}

[†] Laboratory of Molecular Biology and ICRF Unit of Structural Molecular Biology, Department of Crystallography, Birkbeck College, London WC1E 7HX, UK

[§] Department of Molecular Genetics and Protein Chemistry and ^{||} Department of Medicinal Chemistry, Central Research Division, Pfizer Inc., Groton, Connecticut 06340 USA

X-ray analyses have defined the three-dimensional structures of crystals of mouse and human renins complexed with peptide inhibitors at resolutions of 1.9 and 2.8 Å, respectively. The exquisite specificity of renin arises partly from ordered loop regions at the periphery of the binding cleft. Although the pattern of main-chain hydrogen bonding in other aspartic proteinase inhibitor complexes is conserved in renins, differences in the positions of secondary structure elements (particularly helices) also lead to improved specificity in renins for angiotensinogen substrates.

KNOWLEDGE of the three-dimensional structures of proteins is increasingly important in rational approaches to drug design. High-resolution X-ray analyses of enzyme, receptor and other protein complexes can give a detailed description of the general structural features that contribute to specificity. They provide evidence of the importance of individual amino acids and the role of conformational changes and induced-fit recognition of ligands by an enzyme or receptor. This can be used by chemists in their search for a clinically useful candidate.

One important target in recent years has been renin¹. This is the key enzyme of the renin–angiotensin cascade which regulates the formation of angiotensin II, a potent pressor and aldosterogenic substance controlling vascular tone, fluid volume and sodium excretion. Because of the unique specificity of renin, its inhibition is widely expected to provide selective therapy for hypertension, congestive heart failure and associated degenerative disorders linked to angiotensin II. The three-dimensional structures of renin–inhibitor complexes have long been sought as an aid to the discovery of clinically effective antihypertensives².

Sequences of renins^{3,4} show that they belong to the family of aspartic proteinases, several of whose structures are already known⁵. These structures have provided a basis for modelling mouse⁶ and human renins^{7–9}. X-ray analysis of an uncomplexed, partially deglycosylated form of recombinant human renin¹⁰ (atomic coordinates not yet available) has largely confirmed the predictions of the models. But, several loop regions, which are at the periphery of the active site and which may mediate interactions with the substrate, are disordered in the native

human renin crystals. It is not clear whether this is a consequence of the medium resolution data, disorder in the crystal form, deglycosylation of the enzyme or the absence of a bound inhibitor. Although many crystal forms of inhibitor-bound mouse renins have been reported^{11–13}, no high-resolution structures are available.

Structural studies of renin inhibitors complexed with other aspartic proteinases¹⁴ have defined the conformation of the extended main chain of the inhibitor, the location and nature of the specificity subsites and the interactions of various transition state isosteres with the catalytic aspartates of these aspartic proteinases. But many questions remain concerning the specificity of renins for angiotensinogens and the differences in specificities of the human, mouse and other renins. These can be answered unequivocally only by high-resolution X-ray analyses of crystals of inhibitors complexed with renins.

Here we report inhibitor-complexed structures of both mouse submandibular and recombinant human renin. The cubic crystals of human glycosylated renin complexed with a P₄ to P₁' norstatine-containing inhibitor diffract to 2.8 Å resolution and contain two molecules in the asymmetric unit. The crystals of mouse renin complexed with a decapeptide hydroxyethylene isostere inhibitor are monoclinic and contain four molecules in the asymmetric unit. They diffract to a much higher resolution, 1.9 Å. We now describe the analyses of both crystal structures and their refinement to agreement values of 0.19 and 0.18, respectively. The two renins adopt very similar structures in all molecules independently defined in the asymmetric unit; all have well ordered loop structures at the periphery of the active site clefts that differ significantly from those of other aspartic proteinases. There are subtle differences between the subsites of the human and mouse renins that account for their differing

* V.D. and C.G.D. are joint first authors.

† To whom correspondence should be addressed.

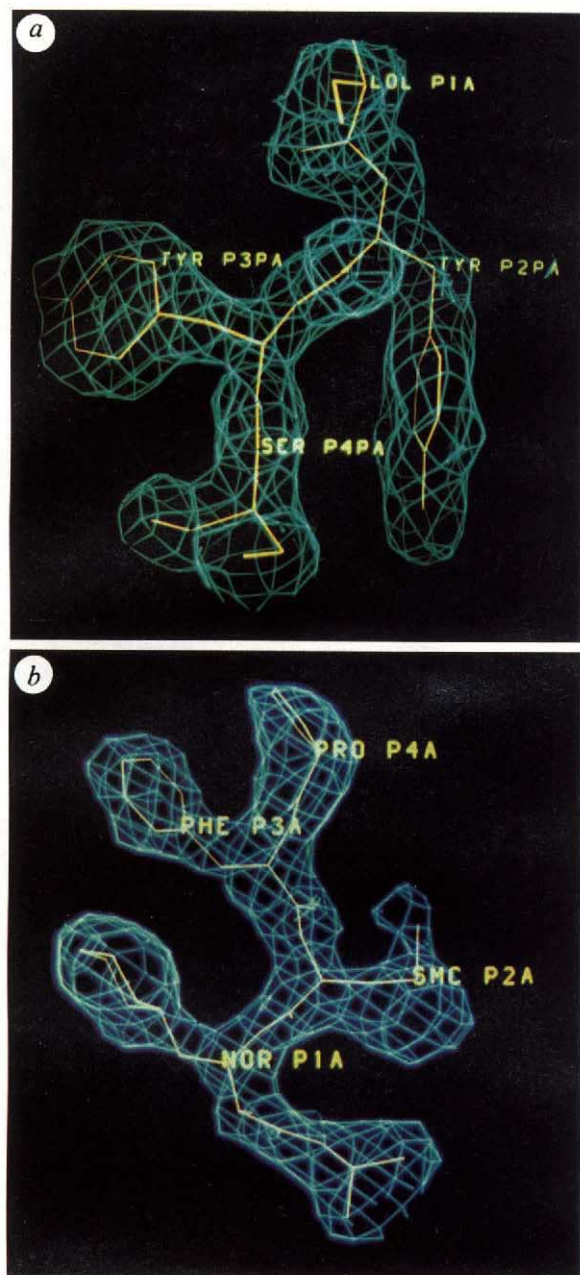


FIG. 1 Electron densities calculated using the Sim-weighted coefficients³³ $2|F_o| - |F_c|$ and displayed on an Evans and Sutherland PS300 graphics terminal using the interactive molecular graphics program FRODO³⁴ of *a*, the mouse and *b*, the human renin inhibitors.

METHODS. The correlation coefficient for the oriented and translated mouse renin model using the parameters listed in Table 1 was 0.34 for reflections in the resolution range 20 to 3 Å. The corresponding *R*-factor was 0.54. Rigid body refinement of the four molecules using RESTRAIN³⁵ corrected the positional parameters for each. The molecules were then split into the three rigid groups defined by Šali *et al.*¹⁸. Further refinement was done using the symmetry-averaging procedure³⁵ treating the secondary structural elements (defined after Kabsch and Sander³⁶) as rigid bodies. After numerous cycles of refinement and manual corrections the individual atomic and thermal parameters of the four molecules and those of 550 solvent molecules have been refined independently to a correlation coefficient of 0.95 and an *R*-factor of 0.18 using 1.9 Å resolution data. The *R*-factor and correlation coefficient for the human renin search model were 0.46 and 0.37, respectively, for the FAST data in the resolution range 4.0–20.0 Å, after the first three cycles of refinement by the method of least-squares using RESTRAIN, treating the two molecules in the asymmetric unit as independent rigid bodies. Subsequently, the two molecules were each split into three rigid groups and later into more rigid domains based on secondary structure. Because there were no major significant differences between the

specificities especially between *S*₁ and *S*₄. The three-dimensional structures of these complexes provide a detailed understanding of the binding of these renin inhibitors, strengthening the basis for the rational design of effective antihypertensive agents.

X-ray analyses of the renin complexes

The preparation of the two renins, the crystallization of their complexes and the X-ray data collection from the crystals have been described in detail elsewhere¹⁵. The cubic crystals (space group *P*2₁3, *a* = *b* = *c* = 143.1 Å; $\alpha = \beta = \gamma = 90^\circ$) of glycosylated human renin complexed with a *P*₄ to *P*₁' norstatine-containing inhibitor CP-85,339 (isopropyl (2*R*,3*S*)-4-cyclohexyl-2-hydroxy-3-((*L*-propyl-*L*-phenylalanyl-5-methyl-*L*-cysteinyl)-amino)butanoate hydrochloride) are isomorphous with crystals of uncomplexed glycosylated human renin¹⁶ and contain two molecules in the asymmetric unit. The crystals of mouse renin complexed with a decapeptide hydroxyethylene isostere inhibitor CH-66 (trimethylacetyl-His-Pro-Phe-His-Leu-ψ(CH(OH)CH₂)-Leu-Tyr-Tyr-Ser-NH₂) are monoclinic (space group *P*2₁, *a* = 78.3 Å, *b* = 117.8 Å, *c* = 85.9 Å; $\alpha = \gamma = 90^\circ$ and $\beta = 101.2^\circ$) and contain four molecules in the asymmetric unit.

The self-rotation function of the human renin complex defines a non-crystallographic 2-fold symmetry that relates the two renin molecules in the asymmetric unit. The self-rotation function of the mouse renin complex indicates pseudo 2-fold axes parallel to *a*^{*} and *c*; thus, four molecules in the asymmetric unit are related by a pseudo orthorhombic (222) symmetry. The structures were solved by molecular replacement using the refined structure of pepsin¹⁷ as search model, and they were refined initially by constrained and then by restrained refinement techniques. Difference electron density maps at intermediate stages revealed the positions of the inhibitors, the saccharide moieties (human renin) and positions of loop structures unique to renins. The current agreement factors and correlation coefficients are 0.19 and 0.91 for human renin at 2.8 Å resolution and 0.18 and 0.95 for mouse renin at 1.9 Å resolution. For human and mouse renins the electron densities in the final calculated maps are of very high quality. Figure 1 shows electron densities for the inhibitors for the mouse and human renins. Neither inhibitor was included in the models used in the molecular replacement.

Multiple copies of the molecules have been independently refined for both renins. Previous structural studies¹⁸ have shown that in different crystalline environments the aspartic proteinases can undergo rigid body movements of a domain comprising residues 190–301 (pepsin numbering) relative to a domain comprising the rest of the molecule, residues 1–189 and 302–326¹⁹. After superposition of the compared enzymes with respect to the first domains, the rigid body movement required to superpose the two resulting orientations and positions of the second domain is described as a screw translocation; this consists of a rotation about and translation along an axis, which passes roughly through the active site aspartates. Comparison of the averaged structure of mouse renin with that of human renin shows that the second domain undergoes a rotation of 1.6° and a translation of 0.4 Å. Very small differences in rigid groups (rotation of ~0.5°; translation of ~0.05 Å) are observed between the independent molecules in the same asymmetric units of human and mouse renins.

two molecules at this stage, refinement was continued using the symmetry-averaging procedure with the Weissenberg imaging plate data. But in the final 15 cycles of refinement, the two molecules were treated as independent copies. After several rounds of refinement punctuated by interactive model-building, the agreement factor and correlation coefficient are 0.19 and 0.91, respectively, for 21,297 reflections; resolution range 8.0 to 2.8 Å. The density maps at this stage clearly indicate the locations of well ordered water molecules. But they have not yet been included in the refinement. The coordinates have been deposited with the Brookhaven Protein Data Bank and will be immediately available. In both X-ray analyses we use a numbering system based on the homology of renins with pepsin.

The human renin is *N*-glycosylated at two sites, but mouse submaxillary renin has no *N*-glycosylation site and it seems that glycosylation is not essential for activity. The first site at Asn 2 in human renin is fully exposed to the solvent and disorder may be expected. The GlcNAc residue attached to Asn 67 is well defined and has been built into the observed density which extends clearly to the second saccharide group.

Overall topology and active site cleft

Figures 2 and 3 show that human and mouse renins have very similar three-dimensional structures, reflecting the high degree of sequence identity (~70%). $C\alpha$ atom pairs superpose with a r.m.s. distance of 0.6 Å, using a 3.5 Å cutoff for pairs of equivalent atoms between the two renin structures. Although the general topology of the two renins is quite similar to that of the known aspartic proteinase structures, the active site cleft has a less open arrangement in renins, partly due to a difference in relative position of the rigid body comprising the C-domain. For instance, there is a rotation of ~4° and translation of ~0.1 Å in the human renin complex with respect to the endothiapepsin-difluorostatone complex²⁰ which makes the cleft smaller in renins.

The entrance to the active site cleft is made even narrower as a consequence of differences in the positions and composition of several well defined loops and secondary structure elements. Unique to the renins is a *cis*-proline, Pro 111, which caps the helix h_{N2} and contributes to the subsites S_3 and S_5 . On an

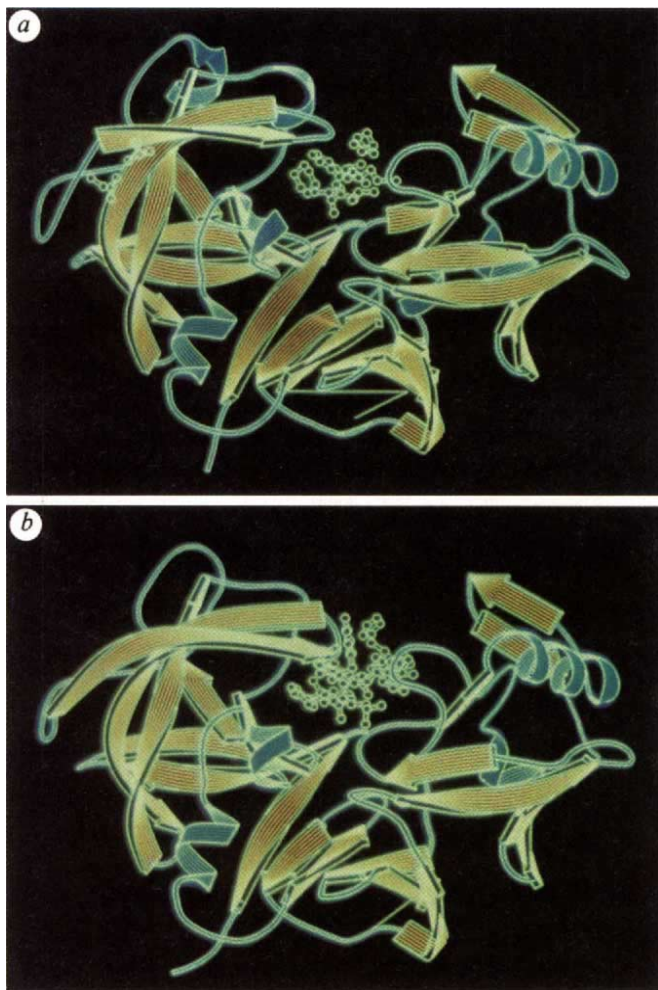


FIG. 2 Schematic cartoons of *a*, human and *b*, mouse renin structures showing the arrangement of α -helices and β -strands. The inhibitors are shown in full.

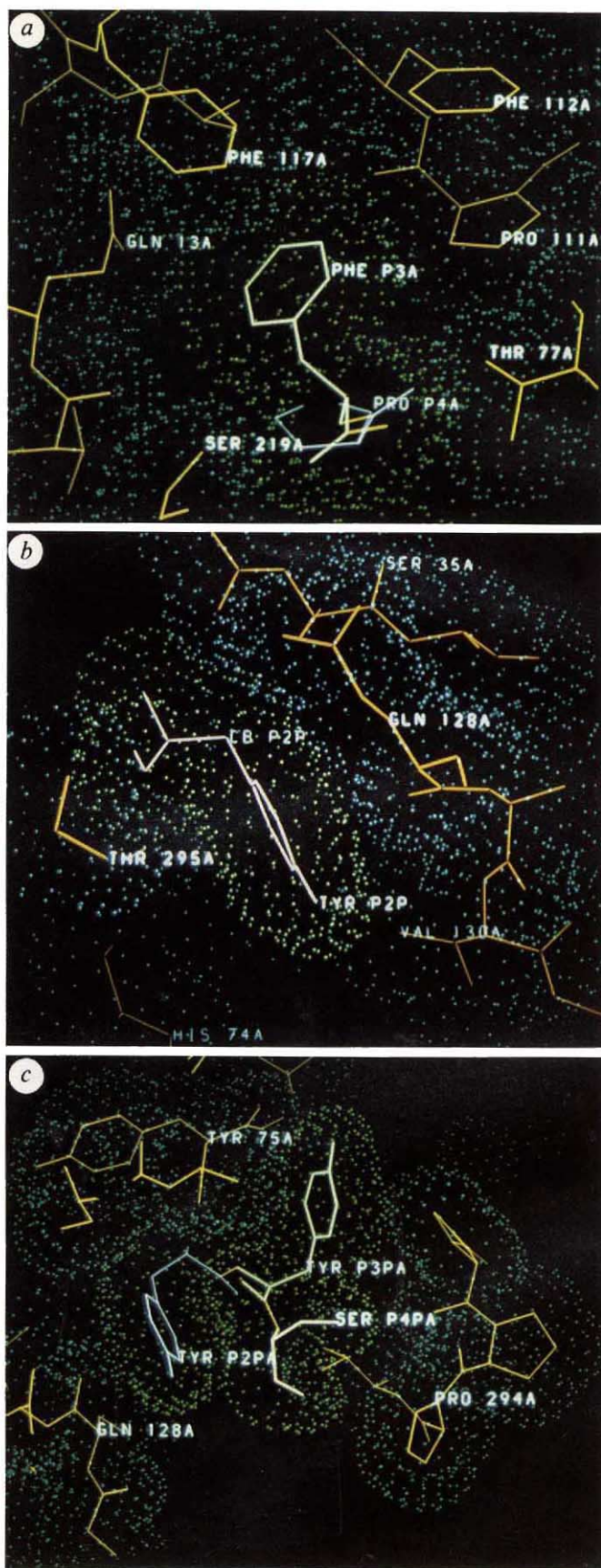


FIG. 5 The specificity pockets in human renin at S_3 (*a*), mouse renin at S'_2 (*b*) and S'_3 and S'_4 (*c*). The bonds between the atoms of the enzymes are drawn in yellow and the van der Waals surface surrounding them are shown as blue-dotted spheres. The surfaces of the corresponding inhibitor atoms are shown surrounded by green-dotted spheres.

equivalent loop in the C-lobe (related by the intra-molecular pseudo 2-fold symmetry), there is a sequence of prolines, the Pro 292-Pro 293-Pro 294 segment, which is also unique to the renins among the aspartic proteinases of known three-dimensional structure, with Pro 294 and Pro 297 in a *cis* configuration. This rather rigid poly-proline loop, and the loop consisting of residues 241–250, lie on either side of the 'flap' formed by residues 72–81. In the inhibitor-bound renins the cleft is covered

by 'flaps' from both lobes rather than from the N lobe alone as in other pepsin-like aspartic proteinases; this gives a superficial similarity to the dimeric, retroviral proteinases where each sub-unit provides an equivalent flap that closes down on top of the inhibitor.

Inhibitor binding and implications for mechanism

Although the human renin inhibitor extends only from P_4 to P_1 , the cyclohexyl norstatine residue at P_1 mimics a dipeptide analogue with its isopropoxy group occupying the subsite for the side-chain of P_1' . Fortunately the mouse renin inhibitor extends further at both ends, from P_6 to P_4' , with all residues clearly defined. Both inhibitors are bound in the extended conformation, similar to that of other aspartic proteinase-inhibitor complexes, and make extensive hydrogen bonds with the enzymes (Fig. 4). The hydrogen bond observed between P_3 N and 219 O γ in other aspartic proteinases is not made in the human renin complex, apparently as a consequence of a repositioning of the protonated Pro P_4 at the N terminus of this short inhibitor.

The locations of the five oxygen atoms of the catalytic aspartyl side chains and the inhibitor hydroxyl group are essentially superimposable in both complexes. The C—OH bonds lie at identical positions when the structures of inhibitor complexes of several aspartic proteinases are superposed, in spite of the differences in the sequence and secondary structure. Most of the complex array of hydrogen bonds found in endothiapepsin complexes can be formed with the exception of that to the threonine or serine at 218, which is replaced by alanine in human renin. This emphasizes the fact that they probably have the same *modus operandi* although their pH optima are quite different.

In the human renin complex, the outer carboxyl oxygen of Asp 215 is hydrogen-bonded to the norstatine ester oxygen. We therefore conclude that it is protonated while Asp 32 is ionized and receives a hydrogen to its inner oxygen from the inhibitor hydroxyl group. This ionization state and hydrogen-bond are those postulated for the stabilization of an endothiapepsin-fluoroketone hydrate complex, as well as the putative tetrahedral intermediate in proteolytic cleavage of the amide bond²⁰.

We have previously suggested that the protonated and buried Asp 303 contributes towards the stabilization of these anionic intermediates⁶ by the formation of a hydrogen-bond between its side-chain carboxylic acid and the main-chain O of Thr 216. In renins this residue is alanine and, as observed by Sielecki *et al.*¹⁰, the carboxylic hydroxyl is replaced by the less acidic phenolic hydroxyl of Tyr 155. This may still contribute to the higher pH profile of human renin. But it seems that this higher pH profile derives also from more general changes in the electrostatic potential which may originate from the substitution of residues more distant from the catalytic residues. As suggested by the experimental observation that human renin cleaves human angiotensinogen with lower K_m and k_{cat} values than those measured with a synthetic human tetradecapeptide substrate²¹, aspects such as the substrate binding and the rate of catalysis may also be influenced to some extent by interactions between the enzyme and its macromolecular substrate at recognition subsites other than those at the active site cleft.

TABLE 1 Coordinates and heights of the peaks in the rotation and translation searches for mouse and human renin inhibitor complexes

Rotation function					
Complex	α	β	γ	Height/r.m.s.	Comments
Human	147.3	22.5	47.3	4.54	Molecule 1
renin	107.0	56.8	356.0	4.00	Molecule 2
Mouse	60.0	20.0	117.5	5.93	
renin	112.5	155.0	290.0	4.26	
Resolution range: 3–20 Å.					
Translation function					
Complex	Orientation	Fractional translations			Height/ r.m.s.
		x	y	z	
Human	Molecule 1	0.419	0.295	0.529	11.8
renin	Molecule 2	0.125	0.410	0.250	10.1
Mouse	Molecule 1	0.382	0.000	0.336	4.9
renin	Molecule 2	0.482	0.536	0.674	8.2
	Molecule 3	0.981	0.028	0.674	7.3
	Molecule 4	0.874	0.565	0.794	10.3

Intensity data from the crystals of mouse renin complex were collected initially on a FAST area detector to 3.0 Å resolution with a completeness of 77% and an R_{merge} of 7.5% and subsequently on the Daresbury Synchrotron Radiation Source to 1.9 Å resolution with a completeness of 86.1% and an R_{merge} of 9.5%. X-ray data from the human renin complex crystals were collected using a FAST electronic area detector to 3 Å resolution with an R_{merge} of 12.9%; these constituted 81% of the total data. Intensity data were also obtained for the same crystal form on a Weissenberg imaging plate using the synchrotron radiation source at the Photon Factory (Tsukuba, Japan) to a resolution of 2.8 Å with an R_{merge} of 7.6% for 22,383 independent reflections which constitute 88.6% of the total data. The coordinates of porcine pepsin¹⁷ were used as the search model in computing the cross-rotation functions of both the complexes using the fast rotation function²⁹ program ALMN³⁰. The rotation function of human renin was complicated and noisy because of the high symmetry of the cubic space group. But the orientations of both the crystallographically independent molecules were derived from distinct peaks. The rotation function of the mouse renin is averaged because of the orientation of the pseudo 2-fold parallel to the unique axis and hence consists of only two peaks as listed above. The orientations of the other two molecules were generated using the pseudo diad parallel to *b*. Because both complexes have more than one copy of the protein per asymmetric unit, the position of one molecule was determined initially by the Crowther and Blow translation function³¹ program TFSGEN³² which uses the interatomic vectors occurring between pairs of crystallographically related molecules. The positions of the others were then derived using the partial structure factor translation function program TFPART which takes into account the vectors between the partially determined structure and molecules related to it by non-crystallographic symmetry³².



FIG. 3 Stereo Ca plot for the comparison of human (solid line) and mouse renin (dashed line) structures. The human renin inhibitor is shown in full.

Specificity

We must now consider how the differences in the specificities of aspartic proteinases are achieved. In modelling exercises^{6-9,22} it was assumed that specificities derive from differences in the sizes of the residues in the specificity pockets (S_n) which complement the corresponding side chains at positions P_n in the substrate/inhibitor²³. But we now show that this simple assumption only partly accounts for the steric basis of specificity.

To compare the binding of the inhibitors in the renin complexes with each other and with other aspartic proteinase complexes, we superposed their three-dimensional structures as a whole and also locally around each of the residues of the inhibitors. This allowed an estimate of the local movements of equivalent residues in the different enzymes relative to the inhibitors. We also listed the enzyme residues lining the specificity pockets in the human and mouse renin complexes (Table 2).

Let us consider first the specificity subsite S_3 . The phenyl rings of Phe $_3$ occupy almost identical positions in both renin inhibitor complexes. Modelling studies have predicted the specificity subsite S_3 to be larger in renins than in other aspartic proteinases²² because of substitution of smaller residues, Pro 111, Leu 114 and Ala 115, in place of larger ones in mammalian and fungal proteinases. But a compensatory movement (~ 5 Å) of the helix h_{N2} makes the pocket quite compact and complementary to the aromatic ring (Fig. 5a). Thus the positions of an element of secondary structure differ between renin and

other aspartic proteinases with a consequent important difference in the specificity pocket.

Differences in the positions of secondary structural elements may also account for the specificity differences at P'_2 . Mouse submaxillary and other non-primate renins do not appreciably cleave human angiotensinogen or its analogues²⁴ which have an isoleucine at P'_2 , although they do cleave substrates with a valine at this position. In contrast, human renin not only cleaves the human and non-primate substrates but also the rat angiotensinogen with tyrosine at P'_2 , albeit rather slowly²¹. This can be explained in terms of the three-dimensional structures. In the mouse renin complex, the P'_2 tyrosyl ring is packed parallel to the adjacent helix h_3 in a narrow pocket (Fig. 5b) and there is only limited space available beyond the $C\beta$ methylene group. This is able to accommodate a valine, but not the larger isoleucine at P'_2 . By contrast, in the human renin, differences in the orientation and position of helix h_3 bring it closer (by ~ 0.5 Å) to the substrate-binding site than in mouse renin. It is orientated in such a fashion in human renin that, although it can accommodate the isobutyl side chain of isoleucine at P'_2 , aromatic rings on substituents such as phenylalanine and tyrosine will have severe short contacts with the side chain of Ile 130 (valine in mouse renin). Thus the reorientation of a helix, coupled with subtle differences in the shapes of the side chains, makes significant changes in the substrate specificity at this subsite.

The three-dimensional structures show that the situation at

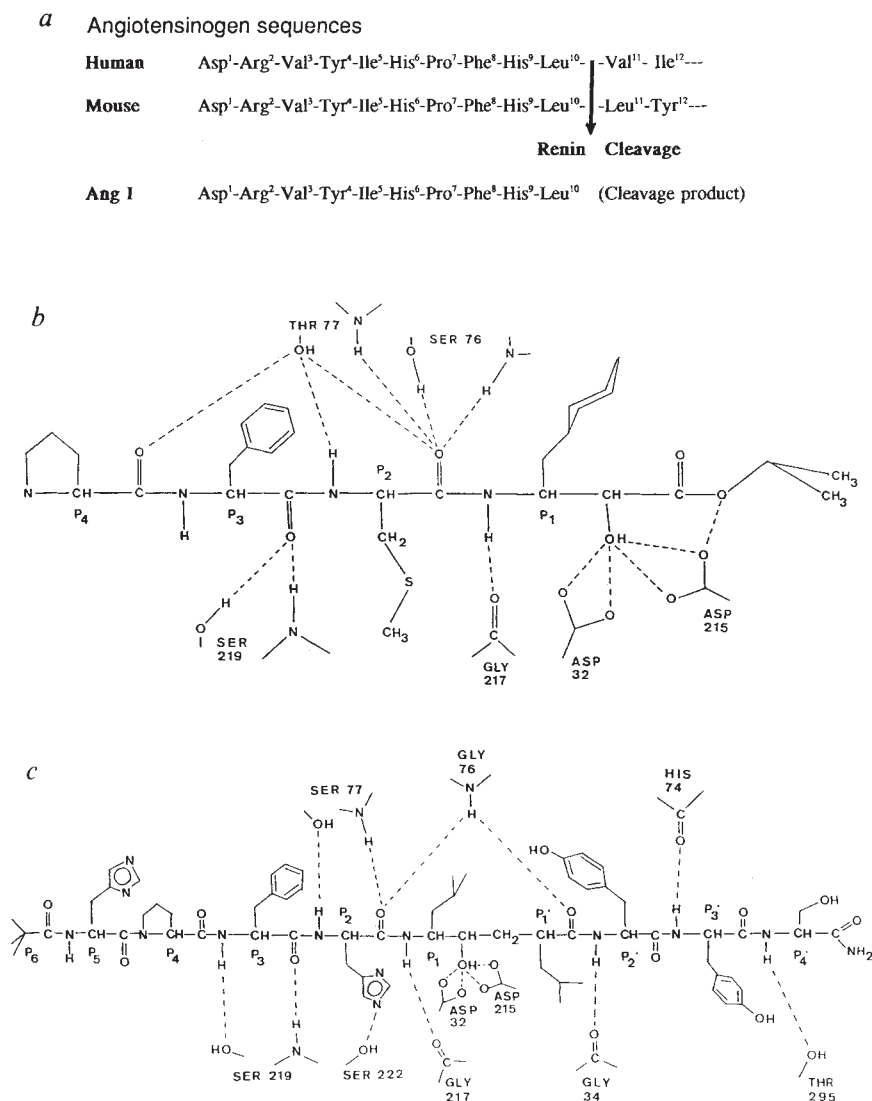


FIG. 4 a, The first 12 residues of the human and mouse angiotensinogens, b, schematic diagrams showing the hydrogen-bonds between human renin and CP-85339 and c, mouse renin and CH-66.

TABLE 2 Specificity pockets for the human and mouse renin complexes

Residues in contact* with the inhibitor		
Subsite	Human renin complex	Mouse renin complex
S6	—	S 12, F 220
S5	—	P 111, L 114
S4	T 77, L 114	S 219, F 220, H 287
S3	Q 13, T 77, P 111, L 114, A 115, F 117, A 218, S 219	Q 13, S 77, P 111, L 114, A 115, F 117, G 217, S 218, S 219
S2	S 76, T 77, G 217, S 222, M 289	Y 75, G 76, S 77, S 218, S 222, H 287, M 289, V 300
S1	D 32, Y 75, F 112, F 117, V 120, D 215, G 217	D 32, G 34, Y 75, V 120, G 217
S1'	G 34, Y 75, S 76, L 213, D 215, I 291	G 34, Y 75, G 76, V 213, D 215, V 298
S2'	—	G 34, S 35, I 73, H 74, Q 128, V 130, T 295
S3'	—	H 74, Y 75, I 291, P 292, T 295
S4'	—	P 294, T 295

* Distance cut-off: 4.0 Å; single-letter amino-acid code.

S₁' is more complex. The specificity subsites at S₁' for the human and mouse renins are complementary to the residues at P₁', which are valine and leucine in human and mouse angiotensinogens^{6,7}, respectively; residue 213 is leucine in human renin and valine in mouse renin. The S₁' pockets of chymosin, pepsin and endothiapepsin have an aromatic side-chain at residue 189 which is replaced in renins by amino acids with smaller side chains (valine in human and serine in mouse renins). This would be expected to make the pocket larger in renins. But X-ray analysis of the mouse renin complex shows that the substrate moves closer to the enzyme in renins as a result of the smaller residue at 189 and the pocket is made even more compact because of changes in the position and composition of the poly-proline loop (residues 290–297). Thus, the specificity difference at this site arises not only from the movement of a secondary structure, in this case a loop region, but also from the substitution of an enzyme residue that allows the substrate to come closer to the body of the enzyme.

The specificity pockets are also influenced by the elaboration of loops on the periphery of the binding cleft in renins. This is most marked at P₃' and P₄'. For instance, in endothiapepsin, which has been the subject of the greatest number of studies, different conformations are adopted at P₃' and the residue at P₄' is generally disordered. In contrast these residues are clearly defined in mouse renin. This is mainly a consequence of the poly-proline loop, illustrated in Fig. 5c. The X-ray analysis of the mouse renin complex shows that the S₃' and S₄' subsites are formed by the poly-proline loop together with residues of the 'flap' and a similar situation is likely to occur in human renin. The well defined interactions of P₃' described in the mouse renin complex explains the significant affinity when inhibitors have phenylalanine or tyrosine at P₃' as well as the importance of a P₃' residue for catalytic cleavage of a substrate by renin²⁵.

In most specificity pockets, hydrogen-bonds between the side chains of the inhibitor and the enzyme do not play a role. But S₂ is an exception. The carbonyl oxygen of P₂ in the human renin complex accepts a hydrogen from the O_γ of Ser 76, which is unique to human renin; residue 76 is a highly conserved glycine in all the other aspartic proteinases, including mouse renin. In mouse renin the P₂ histidyl group forms a hydrogen-bond with the O_γ of Ser 222. If such a conformation were adopted by the substrate in complex with human renin, the two imidazole nitrogens of P₂ histidine would be hydrogen-bonded

to the O_γ of both Ser 76 and Ser 222. This may explain why 'the catalytic efficiency and maximal velocity of human renin are greater above pH 6.5 with the substrate containing histidine at P₂, while they are greater below pH 5.0 with glutamine²⁶ at this position'. The observed reduction in recombinant human renin cleavage rate of a human angiotensinogen analogue containing a 3-methyl histidine substituent at P₂²⁷ could also be explained on the basis of the hydrogen-bonding scheme proposed above.

Implications for rational drug design

In the rational design of drug molecules that are high-affinity inhibitors of enzymes, ligands can be synthesized that emulate the transition state and/or are complementary to the subsites that are sometimes distant from the catalytic site. The three-dimensional structures of the renin-inhibitor complexes described here, together with a further human renin complex²⁸, provide insights into both of these important factors.

With respect to the transition state isostere, the structures of complexes of different aspartic proteinases in the region of the catalytic aspartates and the scissile bond surrogate are virtually identical. This indicates that the catalytic site and mechanism have been highly conserved even in a very divergent family of enzymes that includes the retroviral proteinases and the pepsin-like enzymes. Thus, although the structure provides a blueprint for the design of better transition state isosteres, ligands containing such groups would probably bind well to all aspartic proteinases; we can anticipate that increased affinity would be gained at the expense of decreased specificity.

But subtle differences between aspartic proteinases in the subsites that bind the amino-acid side chains of the substrate provide a basis for the design of ligands that have not only higher affinity but also greater specificity. The studies of the renin-inhibitor complexes reported here confirm the general structural features thought to contribute to renin's specificity²² but demonstrate the need for careful, high-resolution X-ray analyses for more confidence in drug design. In particular, they show that even minor alterations in the positions of secondary structural elements can lead to major changes in the disposition of the subsites. Because these define the species-specificity and determine the catalytic efficiency of the enzymes, a thorough understanding of their composition is indispensable for the synthesis of suitable inhibitors. The specificity pockets are modified by elaboration, particularly of surface loops, which can be disordered in the uncomplexed enzymes and difficult to model with precision from homologous structures.

These data not only establish a new foundation for the rational design of renin inhibitors, but also provide insights into the factors that have general value where protein structures are used to assist the design of clinically useful molecules. □

Received 30 December 1991; accepted 18 May 1992.

1. Tigerstedt, R. & Bergman, P. G. *Skand. Arch. Physiol.* **8**, 223–271 (1898).
2. Greenlee, W. J. *Med. Res. Rev.* **10**, 173–236 (1990).
3. Hobart, P. M., Fogliano, M., O'Connor, B. A., Schaeffer, I. M. & Chirgwin, J. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5026–5030 (1984).
4. Imai, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7405–7409 (1983).
5. Newman, M. et al. *J. molec. Biol.* **221**, 1295–1309 (1991).
6. Blundell, T. L., Sibanda, B. L. & Pearl, L. H. *Nature* **304**, 273–275 (1983).
7. Sibanda, B. L. et al. *FEBS Lett.* **174**, 102–111 (1984).
8. Akahane, K. et al. *Hypertension* **7**, 3–12 (1985).
9. Carlson, W., Karplus, M. & Haber, E. *Hypertension* **7**, 13–20 (1985).
10. Sielecki, A. R. et al. *Science* **241**, 1346–1351 (1989).
11. Mornon, J. P., Surcouf, E., Berthou, J., Covori, P. & Foote, S. J. *J. molec. Biol.* **155**, 539–540 (1982).
12. Navia, M. A., Springer, J. P., Poe, M., Boger, J. & Hoogsteen, K. *J. biol. Chem.* **259**, 12714–12717 (1984).
13. Sibanda, B. L. thesis, Univ. London (1986).
14. Veerapandian, B., Cooper, J. B., Sali, A. & Blundell, T. L. *J. molec. Biol.* **216**, 1017–1029 (1990).
15. Badasso, M. et al. *J. molec. Biol.* **223**, 447–453 (1992).
16. Lim, L. W., Stegeman, R. A., Leimgruber, N. K., Gierse, J. K. & Abdel-Meguid, S. S. *J. molec. Biol.* **210**, 239–240 (1989).
17. Cooper, J. B., Khan, G., Taylor, G., Tickle, I. J. & Blundell, T. L. *J. molec. Biol.* **214**, 199–222 (1990).
18. Sali, A. et al. *EMBO J.* **8**, 2179–2188 (1989).
19. Sali, A. et al. *Proteins* **12**, 158–170 (1992).
20. Veerapandian, B. et al. *Prot. Sci.* **1**, 322–328 (1992).
21. Cumin, F., Le-Nguyen, D., Castro, B., Menard, J. & Corvol, P. *Biochim. biophys. Acta* **913**, 10–19 (1987).

22. Hutchins, C. & Greer, J. *Crit. Rev. Biochem. molec. Biol.* **26**, 77–127 (1991).
23. Schechter, I. & Berger, A. *Biochem. biophys. Res. Commun.* **27**, 157–162 (1967).
24. Poe, M. *et al. Analyt. Biochem.* **140**, 459–467 (1984).
25. Skeggs, L. T., Lentz, K. E., Kahn, J. R. & Hochstrasser, H. *J. exp. Med.* **120**, 13–34 (1968).
26. Green, D. W., Aykent, S., Gierse, J. K. & Zupiec, M. E. *Biochemistry* **29**, 3126–3133 (1990).
27. Holzman, T. F. *et al. J. Prot. Chem.* **10**, 553–563 (1991).
28. Rahuel, J., Priestle, J. P. & Grütter, M. G. *J. struct. Biol.* **107**, 227–236 (1991).
29. Crowther, R. A. in *The Molecular Replacement Method* (ed. Rossman, M. G.) 173–178 (Gordon and Breach, New York, 1972).
30. Dodson, E. J. in *Molecular Replacement, Proceedings of the Daresbury Study Weekend* (ed. Machin, P.) 33–45 (Daresbury Laboratory, DL/SCI/R23, 1985).
31. Crowther, R. A. & Blow, D. M. *Acta Crystallogr.* **23**, 544–548 (1967).
32. Driessen, H. P. C. *et al. Acta Crystallogr. B* **47**, 987–997 (1992).

33. Read, R. J. *Acta Crystallogr. A* **42**, 140–149 (1986).
34. Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
35. Driessen, H. P. C. *et al. J. appl. Crystallogr.* **22**, 510–516 (1989).
36. Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).

ACKNOWLEDGEMENTS. We thank M. Szelke, J. Sueiras-Diaz and D. M. Jones for supplying us with the mouse renin inhibitor, K. Murakami and H. Miyazaki for supplying the samples of mouse renin that were used in crystallization trials, D. B. Damon (Pfizer Inc.) for synthesizing the human renin inhibitor CP-85,339 B. W. Dominy (Pfizer Inc.) for discussions and modelling assistance, R. Sarra (Birkbeck College) for helpful discussions and N. Sakabe and A. Nakagawa for their assistance at the Photon Factory, KEK Laboratory, Tsukuba in collection of the human renin X-ray data. We thank the AFRC (M.N. and V.D.), SERC (C.A., C.D. and M.B.) and ICRF (J.B.C., H.P.C.D.) for financial support.

LETTERS TO NATURE

Millisecond pulsars with extremely strong magnetic fields as a cosmological source of γ -ray bursts

V. V. Usov

Physics Department, Weizmann Institute of Science, Rehovot 76100, Israel

THE spatial and luminosity distribution of γ -ray bursts as observed by the BATSE instrument on the Compton Gamma Ray Observatory^{1,2} provides support for the revival of the idea^{3,4} that the burst sources are at cosmological distances⁵. I present here a new model for γ -ray bursts at cosmological distances, based on the formation of rapidly rotating neutron stars with surface magnetic fields of the order of 10^{15} G. Such objects could form by the gravitational collapse of accreting white dwarfs with anomalously high magnetic fields in binaries, as in magnetic cataclysmic binaries. Once formed, such rapidly rotating and strongly magnetized neutron stars would lose their rotational kinetic energy catastrophically, on a timescale of seconds or less: rotation of the magnetic field creates a strong electric field, and hence an electron–positron plasma, which I show to be optically thick and in quasi-thermodynamic equilibrium. This plasma flows away from the neutron star at relativistic speeds, and X-ray and γ -ray emission at the photosphere of this relativistic wind may then reproduce the observational characteristics of a γ -ray burst.

Let us consider a close binary system in which one component is a white dwarf with a surface magnetic field, $\tilde{B}_s$, of a few times 10^9 G. The secondary star might be a red or white dwarf which fills its inner critical Roche lobe. The strength of the magnetic field is known only for a few tens of white dwarfs which are not far from the Sun; for these, B_s is in the range $\sim 10^5$ G to $\sim 10^9$ G (refs 6, 7). It is therefore reasonable to assume that magnetic white dwarfs exist whose surface magnetic field is a few times more than the maximum observed values.

The predecessor of such a binary may form by tidal capture in dense stellar systems, such as globular clusters⁸. If the orbital period of the created binary is smaller than a few hours, the characteristic time over which the orbital period varies because of gravitational radiation is less than the age of the Universe. During evolution, the components of the binary come closer together because gravitational radiation is generated. At some point, the companion of the white dwarf fills its Roche lobe and the binary is formed. Material transfers to the magnetic white dwarf from its companion. When the mass of the accreting magnetic white dwarf increases up to the Chandrasekhar mass, it collapses, and a neutron star forms.

In the process of collapse, the stellar radius changes from $\tilde{R} \approx 10^9$ cm for the white dwarf to $R \approx 10^6$ cm for the neutron star. From magnetic-flux conservation it follows that the strength

of the magnetic field on the surface of the formed neutron star can reach $B_s = \tilde{B}_s (\tilde{R}/R)^2$, which is a few times 10^{15} G. This value of B_s is two orders of magnitude higher than the maximum magnetic fields observed for pulsars⁹.

The typical angular velocity of a binary consisting of two dwarf stars in the state of mass transfer from one to the other is $\tilde{\Omega} \approx 10^{-3} - 10^{-1} \text{ s}^{-1}$. The rotation of a white dwarf with a very strong magnetic field is synchronized with the orbital rotation by the electromagnetic interaction between the components of the binary¹⁰. From angular-momentum conservation during gravitational collapse, the expected angular velocity of the created neutron star is $\Omega \approx \tilde{\Omega} (\tilde{R}/R)^2 \approx 10^4 \text{ s}^{-1}$ for $\tilde{\Omega} \approx 10^{-2} \text{ s}^{-1}$. This value of Ω is near the maximum possible value for a neutron star¹¹. A very rapidly rotating neutron star, resembling known millisecond pulsars, with an unusually strong magnetic field may therefore be created on gravitational collapse of a magnetic white dwarf in a close binary.

The rotational kinetic energy of a neutron star with moment of inertia $I \approx 10^{45} \text{ g cm}^2$ is

$$E_{\text{kin}} = \frac{1}{2} I \Omega^2 \approx 5 \times 10^{52} \text{ erg for } \Omega \approx 10^4 \text{ s}^{-1} \quad (1)$$

which is enough to explain the total energy released during the burst if the γ -ray bursts are of cosmological origin^{3–5}.

The rotation of the magnetic neutron star decelerates because of the electromagnetic torque and the torque connected with the generation of gravitational-quadrupole radiation^{12–15}. For the first mechanism, the rate of kinetic energy loss is of the order of the magnetic-dipole luminosity of a magnetic neutron star rotating in vacuum^{12–15}

$$-\frac{dE_{\text{kin}}}{dt} \approx L_{\text{md}} = \frac{2}{3} \frac{\mu^2 \Omega^4}{c^3} \approx 2.2 \times 10^{51} \left(\frac{B_s}{3 \times 10^{15} \text{ G}} \right)^2 \left(\frac{\Omega}{10^4 \text{ s}^{-1}} \right)^4 \text{ erg s}^{-1} \quad (2)$$

where $\mu = B_s R^3$ is the magnetic-dipole moment of the neutron star.

The luminosity of the rotating neutron star in gravitational-quadrupole radiation is¹³

$$L_{\text{gw}} = \frac{32}{5} \frac{G \varepsilon^2 I^2 \Omega^6}{c^5} \approx 1.8 \times 10^{55} \varepsilon^2 \left(\frac{I}{10^{45} \text{ g cm}^2} \right)^2 \left(\frac{\Omega}{10^4 \text{ s}^{-1}} \right)^6 \text{ erg s}^{-1} \quad (3)$$

Here G is the gravitation constant and ε is the equatorial ellipticity of a neutron star.

If the magnetic field is responsible for the equatorial ellipticity, we have¹³

$$\varepsilon \approx \frac{25}{24} \frac{R^4}{GM^2} (3B_p^2 - \langle B_\phi^2 \rangle) \approx 10^{-4} \frac{B_p^2 - \frac{1}{3} \langle B_\phi^2 \rangle}{(3 \times 10^{15} \text{ G})^2} \quad (4)$$

where B_p and B_ϕ are poloidal and toroidal components of magnetic field. In this case the kinetic energy losses through gravitational-quadrupole radiation are not essential for the

deceleration of the rotation if the strength of the magnetic field inside the neutron star is $\leq 3 \times 10^{17}$ G.

The ellipticity and gravitational-quadrupole radiation increase sharply because of the neutron star's instability when its angular velocity is larger than some critical value Ω_{cr} (ref. 11). The critical angular velocity depends on mass and is $\sim (0.4-1.2) \times 10^4 \text{ s}^{-1}$. If $\Omega > \Omega_{cr}$ the values of ε and L_{gw} are a few times 0.1 and $\sim 10^{52}-10^{53} \text{ erg s}^{-1}$, respectively^{11,16}.

The characteristic time for kinetic energy release is

$$\tau = \frac{E_{kin}}{L_{md} + L_{gw}} \approx \begin{cases} 20(B_s/3 \times 10^{15} \text{ G})^{-2}(\Omega/10^4 \text{ s}^{-1})^{-2} \text{ s} & \text{if } \Omega < \Omega_{cr} \\ 3 \times 10^{-3} \varepsilon^{-2}(\Omega/10^4 \text{ s}^{-1})^{-4} \text{ s} & \text{if } \Omega \geq \Omega_{cr} \end{cases} \quad (5)$$

The energy of rotation released because of the action of the electromagnetic torque is transferred into electromagnetic radiation, a relativistic electron-positron plasma and ultra-relativistic high-energy particles by means of the following sequence of processes^{14,17,18}: (1) generation of a strong electric field near the rotating magnetic neutron star; (2) acceleration of primary particles by a strong electric field to ultra-relativistic energies; (3) generation of γ -rays by the accelerated particles; (4) creation of a secondary relativistic electron-positron plasma by the absorption of γ -ray quanta in a strong magnetic field ($\gamma + B \rightarrow e^+ + e^- + B$); (5) generation of radiation by the electron-positron plasma. This sequence of processes also holds for millisecond pulsars with superstrong magnetic fields, but some physical processes are changed qualitatively.

The component of electric field $E_{\parallel}$ along magnetic field B in the magnetosphere of a young pulsar may be as high as¹⁴

$$E_{\parallel}^{\max} = \frac{\mathbf{E}^{\max} \cdot \mathbf{B}}{|\mathbf{B}|} \approx B_s \left(\frac{\Omega R}{c} \right)^{5/2} \approx 2 \times 10^{14} \left(\frac{B_s}{3 \times 10^{15} \text{ G}} \right) \left(\frac{\Omega}{10^4 \text{ s}^{-1}} \right)^{5/2} \text{ e.s.u.} \quad (6)$$

For a millisecond pulsar with superstrong magnetic field we have $E_{\parallel}^{\max} > E_{cr} = m^2 c^3 / e \hbar = 4.41 \times 10^{13} \text{ e.s.u.}$ (here $\hbar$ is Planck's constant, m and e are the mass and charge of the electron, respectively). A vacuum with such a strong electric field is unstable¹⁹ and electron-positron pairs are created directly ($E \rightarrow e^+ + e^- + E$) without any generation of γ -ray quanta subsequent absorption. This mechanism may be essential in the magnetosphere of millisecond pulsars with extremely strong magnetic fields in addition to the electron-positron pair creation through the processes $\gamma + B \rightarrow e^+ + e^- + B$ and $\gamma' + \gamma'' \rightarrow e^+ + e^-$.

If $E_{\parallel} > 0.1 E_{cr}$, the rate of direct pair production by the electric field is extraordinarily powerful¹⁹. Therefore, very strong electric fields, such as $E_{\parallel}^{\max} \approx 10^{14} \text{ e.s.u.}$, cannot exist in pulsar magnetospheres and are screened by the created pairs up to $E_{\parallel} \leq 0.1 E_{cr}$.

The density of electron-positron pairs near the light cylinder of a millisecond pulsar with an extremely strong magnetic field may be estimated as

$$n_p \approx \frac{L_p}{4\pi r_{lc}^2 c \langle \Gamma_p \rangle m c^2} \approx 4 \times 10^{32} \alpha \left(\frac{B_s}{3 \times 10^{15} \text{ G}} \right)^2 \left(\frac{\Omega}{10^4 \text{ s}^{-1}} \right)^6 \langle \Gamma_p \rangle^{-1} \text{ cm}^{-3} \quad (7)$$

where L_p is the luminosity of the pulsar in electron-positron pairs, $\alpha = L_p / L_{md}$,

$$r_{lc} = c / \Omega = 3 \times 10^6 (\Omega / 10^4 \text{ s}^{-1}) \text{ cm} \quad (8)$$

is the radius of the pulsar light cylinder and $\langle \Gamma_p \rangle$ is the mean Lorentz factor of created electron-positron pairs ($\Gamma = (1 - v^2/c^2)^{-1/2}$, where v is the particle velocity).

The optical depth for electron-positron pair annihilation near the light cylinder is roughly

$$\tau_{ann} \approx r_{lc} n_p \sigma_{ann} \approx 10^{15} \alpha \left(\frac{B_s}{3 \times 10^{15} \text{ G}} \right)^2 \left(\frac{\Omega}{10^4 \text{ s}^{-1}} \right)^5 \langle \Gamma_p \rangle^{-2} \quad (9)$$

where σ_{ann} is the cross-section of electron-positron pair annihilation²⁰.

Let us estimate the values of α , $\langle \Gamma_p \rangle$ and τ_{ann} . The observed luminosities of young pulsars, such as PSR0531+21, PSR0540-69 and PSR0833-45, in electromagnetic radiation and in very-high-energy particles are $\sim 10^{-3} L_{md}$ and a few times $0.1 L_{md}$, respectively. The expected luminosity, L_p , in electron-positron pairs is somewhere between these two values. As was mentioned above, the component of electric field $E_{\parallel}$ along the magnetic field is very strong in the magnetosphere of a millisecond pulsar with an extremely strong magnetic field. In this case most of the energy of ultra-relativistic primary particles is lost as γ -radiation and absorbed near the pulsar with electron-positron pair creation¹⁷. Therefore the luminosity of such a pulsar in electron-positron pairs is most probably a few times $0.1 L_{md}$, that is the value of α is a few times 0.1. The mean value of the Lorentz factor of electron-positron pairs outflowing from the environment of ordinary pulsars is about 10 (refs 14, 21). In a magnetic field $B \geq 10^{13} \text{ G}$, the decay of γ -rays ($\gamma + B \rightarrow \gamma' + \gamma'' + B$) is essential^{22,23}. The result of this process is some decrease of the mean Lorentz factor of outflowing pairs, so in our case the expected value of $\langle \Gamma_p \rangle$ is less than 10.

Using these values of α and $\langle \Gamma_p \rangle$, we can see from equation (9) that near the light cylinder of the pulsar, with B_s a few times 10^{15} G and $\Omega \approx 10^4 \text{ s}^{-1}$, the optical depth for electron-positron pair annihilation is very large ($\tau_{ann} \approx 10^{13}-10^{14} \gg 1$). The optical depths for Compton scattering²⁴, and for γ -ray absorption²⁵⁻²⁷ ($\gamma + B \rightarrow e^+ + e^- + B$ and $\gamma + \gamma' \rightarrow e^+ + e^-$) and photon decay^{22,23} ($\gamma + B \rightarrow \gamma' + \gamma'' + B$), are also very large near the pulsar light cylinder as well. Therefore the electron-positron plasma and radiation must be in quasi-thermodynamical equilibrium in the environment of a millisecond pulsar with a very strong magnetic field. This is one of the main differences between such a pulsar and ordinary known pulsars.

The electron-positron plasma flows away from the pulsar with Lorentz factor $\langle \Gamma_p \rangle$ between 1 and 10. During outflow, the electron-positron plasma accelerates and its density decreases^{28,29}. At a distance r_{rad} from the pulsar where the optical depth for X-rays and γ -rays is ~ 1 , the radiation propagates freely. From estimates (V.V.U., manuscript in preparation) it follows that the radius of the γ -ray photosphere r_{rad} for an electron-positron wind outflowing from a millisecond pulsar with extremely strong magnetic field is roughly the white dwarf's radius, $r_{rad} \approx 10^8-10^9 \text{ cm}$. The mean Lorentz factor of plasma particles at $r \approx r_{rad}$ is $\langle \Gamma_{rad} \rangle \approx r_{rad} / r_{lc}$, which is a few times ($10-10^2$) (refs 5, 30).

The typical energy of γ -rays $\mathcal{E}_{\gamma}$ radiated by the photosphere of the electron-positron wind is a few times (0.1-1) MeV (refs 3, 30-33, and V.V.U., manuscript in preparation). The high-energy tail of the γ -ray spectrum can be up to $\sim \langle \Gamma_{rad} \rangle \text{ MeV}$, or a few hundred MeV. Because $L_{md} \propto t^{-2}$ for $\Omega < \Omega_{cr}$ (ref. 13) the γ -ray flux must decrease with time t on average as t^{-2} .

Now let us consider the explanation of the main observational data on γ -ray bursts. The expected luminosity of our millisecond pulsar is a few times $0.1 L_{md} \approx 10^{51} \text{ erg s}^{-1}$ (see equation (2)). This coincides with the value needed to explain the cosmological γ -ray bursts of which the sources are at redshift ~ 1 (refs 3-5). The typical energy of γ -rays, $\mathcal{E}_{\gamma} \approx (0.1-1) \text{ MeV}$, radiated by the photosphere of the electron-positron wind agrees with γ -ray spectrum observations for the bursts²⁰. Observed durations of γ -ray bursts cover a range from ~ 0.1 to $\sim 10^3 \text{ s}$, with a mean duration $\sim 10-20 \text{ s}$. The characteristic time of kinetic energy release given by equation (5) overlaps the range of the γ -ray burst durations. The present model can explain not only the

durations of the γ -ray bursts but also their light curves. It is expected that the flux of γ -rays rises sharply as a rule. The rise time may be as short as $\sim r_{\text{rad}}/c(\Gamma_{\text{rad}})^2 \approx 10^{-4} - 10^{-3}$ s (ref. 3). Then the behaviour of the γ -ray flux depends essentially on the relation between the angular velocity of the pulsar Ω_0 at the moment of its creation and the value of Ω_{cr} . If $\Omega_0 > \Omega_{\text{cr}}$, the deceleration of the pulsar rotation is determined by the generation of gravitational waves. In this case the characteristic time of the γ -ray flux decay is $\tau \approx 0.01 - 0.1$ s. When the value of Ω becomes less than Ω_{cr} the γ -ray flux decays with time constant ~ 10 s or more. If $\Omega_0 < \Omega_{\text{cr}}$, the phase with small time constant is absent. Most probably, the short γ -ray bursts, of duration less than ~ 1 s and with single peaked time profiles, are generated by the pulsar when $\Omega_0 > \Omega_{\text{cr}}$. For these γ -ray bursts the sensitivity of γ -ray telescopes was insufficient to detect a tail with large time constant generated at $\Omega < \Omega_{\text{cr}}$.

The γ -ray flux from relativistic electron-positron wind may vary on the scale of the rise time, that is on a timescale $\sim 10^{-4} - 10^{-3}$ s. This is short enough to explain the γ -ray flux variations observed for the bursts. The expected period of the neutron star precession is³⁴ $P_{\text{pr}} \approx \varepsilon^{-1} P \approx 1 - 10$ s for $P = 2\pi/\Omega \approx 10^{-3}$ s and $\varepsilon \sim 10^{-3} - 10^{-4}$ (see equation (4)). The value of P_{pr} is in agreement with the periodicity of the γ -ray flux oscillations observed for some bursts²⁰.

The rotational kinetic energy of a millisecond pulsar with an extremely strong magnetic field becomes smaller than 10^{47} erg in a few months after its formation. This energy is too small to be responsible for the subsequent activity of pulsars at cosmological distances. The pulsar magnetic energy $E_m \sim R^3 B_s^2 \sim 10^{49} - 10^{50}$ erg is high enough to be a source of pulsar activity after the γ -ray burst generation. Some of the magnetic energy may be released suddenly because, for example, of the oscillatory instability of a magnetic neutron star³⁵. The development of this instability may be responsible for the recurrent events observed for a few γ -ray bursters.

The creation of a millisecond pulsar with an extremely strong magnetic field is a very rare event in a galaxy. To explain the rate of the γ -ray bursts observed by BATSE, such events must occur at a rate of $\sim 10^{-8} \text{ yr}^{-1}$ per galaxy^{4,5,36}. $\square$

Real-time imaging of the magnetic flux distribution in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$

C. A. Durán, P. L. Gammel, R. Wolfe, V. J. Fratello, D. J. Bishop, J. P. Rice* & D. M. Ginsberg*

AT & T Bell Laboratories, Murray Hill, New Jersey 07974, USA

* Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

THE magnetic response of the high-transition-temperature oxide superconductors to an applied field is related to two technologically important issues in these materials: their finite resistivities well below the superconducting transition and their critical currents. In addition, the scientifically interesting issues of pinning, melting, vortex glass behaviour and vortex liquids¹ can be probed by studying the time dependence of the magnetization. Most techniques used to probe the magnetization of superconductors represent averages over the entire volume of the sample. This precludes the study of defects, and also raises the question of whether the observed behaviour is intrinsic or is dominated by defects. On the other hand, spatially resolved techniques such as magnetic decoration or scanned probe microscopy are usually too slow to allow studies of the dynamics of these systems. Here we demonstrate a technique for real-time, spatially resolved imaging of flux dynamics in superconductors with a wide variety of sample morphologies. Following earlier work²⁻⁶, we place a wafer of a magneto-optical material—in this case, a doped garnet film—in contact with a superconducting sample in a polarizing light microscope. The spatial and temporal resolution afforded by the high-quality garnet film allows us to investigate the effect of defects (specifically, twin planes) on the magnetic response of the superconductor.

Our apparatus is shown schematically in Fig. 1. We use a bismuth- and gallium-doped yttrium-iron-garnet (YIG) film placed close to the superconductor⁴⁻⁶. The magnetic response of the garnet and hence of the superconductor is probed using polarized light in an optical microscope, whose output is recorded with a CCD camera for later computer processing. Much work has gone into developing garnet films for use in magnetic bubble technology⁷; we exploit the availability of the high-quality films developed for this. Our films were grown by liquid phase epitaxy on gadolinium gallium garnet (GGG) substrates. The high quality is essential because the time constant for domain motion ($< 1 \mu\text{s}$) is much shorter than the other time constants in the systems that we wish to study. In addition, the coercivity of the films remains low down to 4 K, allowing us to measure the temperature dependence of dynamical processes in the superconductor over a wide range without having to correct for the response of the garnet itself.

The garnet imaging technique relies on the fact that the (Bi, Ga):YIG films have a strong uniaxial anisotropy and the ferrimagnetic domains are perpendicular to the plane of the film. In zero applied field, there are on average as many domains up as down. When a field is applied parallel to the easy axis of the garnet, the domain structure changes to produce a magnetization which is linear in field at low fields. If the applied magnetic field is inhomogeneous, then the local domain density will reflect that distribution of externally applied field.

The domains in the garnet can be observed using Faraday rotation ($\sim 5^\circ$ for these films) in a polarizing light microscope. Because the superconducting samples are not transparent, we have mirrored one side of the garnet with $\sim 2,000 \text{ \AA}$ of silver. We probe the domains using reflected light as shown in Fig. 1. The silver film is transparent enough to allow us to observe the outline of the sample in transmission so that we can register the domain patterns with the superconductor. This trick allows us to register the patterns of magnetization that we measure to an

Received 20 February; accepted 1 May 1992.

1. Fishman, G. J. et al. *Proc. 2nd Gamma Ray Observatory Science Workshop* (in the press).
2. Meegan, C. A. et al. *Nature* **355**, 143–145 (1992).
3. Priulitski, O. F. & Usov, V. V. *Astrophys. Space Sci.* **34**, 395–401 (1975).
4. Usov, V. V. & Chibisov, G. V. *Soviet Astr.* **19**, 115–116 (1975).
5. Paczynski, B. *Acta astr.* **41**, 257–267 (1991).
6. Schmidt, G. D. & Liebert, J. *Astrophys. Space Sci.* **131**, 549–557 (1987).
7. Latter, W. B., Schmidt, G. D. & Green, R. F. *Astrophys. J.* **320**, 308–314 (1987).
8. Lightman, A. P. & Grindlay, J. E. *Astrophys. J.* **262**, 145–152 (1982).
9. Manchester, R. N. & Taylor, J. H. *Astr. J.* **86**, 1953–1973 (1981).
10. Lamb, F. K., Aly, J.-J., Cook, M. C. & Lamb, D. Q. *Astrophys. J.* **274**, L71–L75 (1983).
11. Friedman, J. L. *Phys. Rev.* **51**, L11–L18 (1983).
12. Pacini, F. *Nature* **219**, 145–147 (1968).
13. Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **157**, 1395–1417 (1969).
14. Arons, J. in *Proc. Workshop Plasma Astrophysics*, ESA SP-161, 273–286 (European Space Agency, Varenna, 1981).
15. Michel, F. C. *Rev. mod. Phys.* **54**, 1–66 (1982).
16. Chandrasekhar, S. *Astrophys. J.* **161**, 571–578 (1970).
17. Sturrock, P. A. *Astrophys. J.* **164**, 529–556 (1971).
18. Ruderman, M. A. & Sutherland, P. G. *Astrophys. J.* **196**, 51–72 (1975).
19. Schwinger, J. *Phys. Rev.* **82**, 664–672 (1951).
20. Harding, A. K. *Phys. Rep.* **206**, 327–391 (1991).
21. Machabelli, G. Z. & Usov, V. V. *Soviet Astr. Lett.* **15**, 393–397 (1989).
22. Adler, S. L. *Ann. Phys.* **67**, 599–647 (1971).
23. Usov, V. V. & Shabad, A. E. *Sov. Astr. Lett.* **9**, 212–214 (1983).
24. Ochelkov, Yu. P. & Usov, V. V. *Astrophys. Space Sci.* **96**, 55–81 (1983).
25. Erber, T. *Rev. mod. Phys.* **38**, 626–647 (1966).
26. Shabad, A. E. & Usov, V. V. *Nature* **295**, 215–217 (1982).
27. Cheng, A. F. & Ruderman, M. A. *Astrophys. J.* **214**, 598–606 (1977).
28. Paczynski, B. *Astrophys. J.* **308**, L43–L46 (1986).
29. Goodman, J. *Astrophys. J.* **308**, L47–L50 (1986).
30. Shemi, A. & Piran, T. *Astrophys. J.* **365**, L55–L58 (1990).
31. Cavallo, G. & Rees, M. J. *Mon. Not. R. astr. Soc.* **183**, 359–365 (1978).
32. Heuter, G. F. & Lingelfelter, R. E. in *Positron-Electron Pairs in Astrophysics* (eds Burns, M. L., Harding, A. K. & Ramaty, R.) 89–93 (American Institute of Physics, New York, 1983).
33. Horstman, H. M. & Cavallo, G. *Astr. Astrophys.* **122**, 119–123 (1983).
34. Shahan, J. J. *Phys.* **41**, C2-9–C2-23 (1980).
35. Usov, V. V. *Astrophys. Space Sci.* **107**, 191–197 (1984).
36. Dermer, C. D. *Phys. Rev. Lett.* **68**, 1799–1802 (1992).

ACKNOWLEDGEMENTS. I thank M. Milgrom for helpful conversations.

optical micrograph of the sample to better than one pixel ($\sim 5 \mu\text{m}$). Thus we can be sure that we are correctly attributing the magnetic response that we see to specific defects or other regions in the crystal.

The superconductor itself is pressed against the garnet with a typical separation of $3\text{--}5 \mu\text{m}$. The sample and garnet are mounted on the cold finger of a helium-flow cryostat with a diode thermometer in close proximity. We record the images at 30 frames s^{-1} using a CCD camera and an optical disk recorder, and subsequently analyse them on a personal computer.

The local field is measured by averaging the grey level over an area which must be several times the size of one domain. With typical domain sizes ranging down to $2 \mu\text{m}$, this gives us a resolution of $\sim 5 \mu\text{m}$. Converting the domain patterns to field intensity is straightforward because the grey level is linear in field over our field range. In practice, we calibrate the grey scale far from the sample where the local field equals the applied field. In addition, spurious effects such as uneven illumination and optical interference patterns can be easily removed from the images during the analysis.

The intensity averaging can also be done optically by choosing a magnification such that a pixel in the final image corresponds to the desired averaging area. In this case the grey level of a single pixel directly gives the value of the local field. This technique allows us to study large areas of the sample (several square millimetres) with the ultimate spatial resolution allowed by the garnet and still remain within the capability of standard video equipment.

Here we present the results of two kinds of measurements of flux dynamics in lightly twinned $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ (YBCO) single crystals⁸. In the first type, we study the magnetization in an applied oscillating magnetic field. We can study the response from 100 Hz to 10^{-5} Hz and extract the amplitude and phase of the various Fourier components as a function of field, position and temperature. In principle the response in our system could be measured up to 1 MHz with different recording equipment. Nonlinear processes are measured and can be studied by extracting the higher-order Fourier components of the response (that is, at multiples of the applied field frequency).

Figure 2a is a polarized light micrograph of our lightly twinned YBCO sample. The light and dark areas are the regions in which orientations of the *a*- and *b*-axes differ by 90° . Figure 2b shows the real part of the Fourier component of the response of the crystal at the fundamental frequency with an applied field

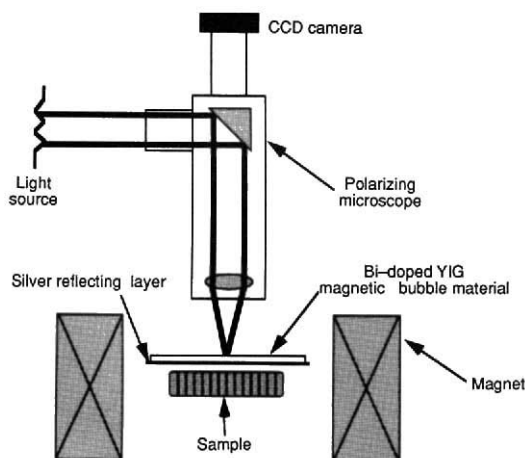


FIG. 1 Schematic diagram of the experimental arrangement. The magnetic garnet film touches the sample, which is attached to the cold finger of a liquid helium-flow cryostat. The CCD video camera is connected to an optical disk recorder and a computer, which allows real-time data acquisition, and frame-by-frame image digitization and analysis.

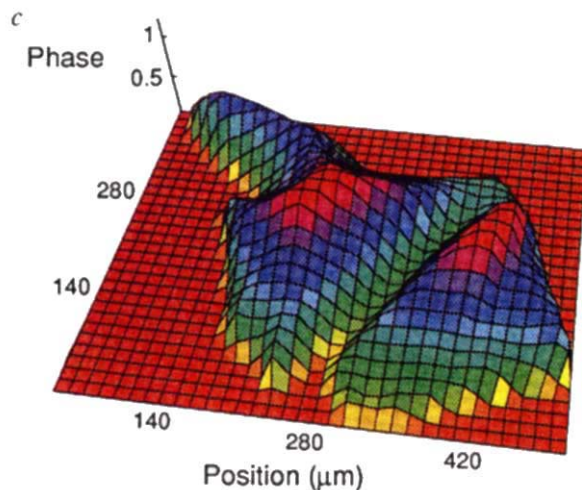
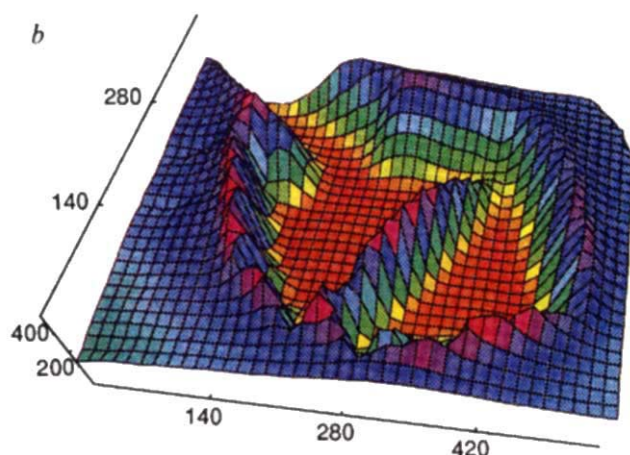
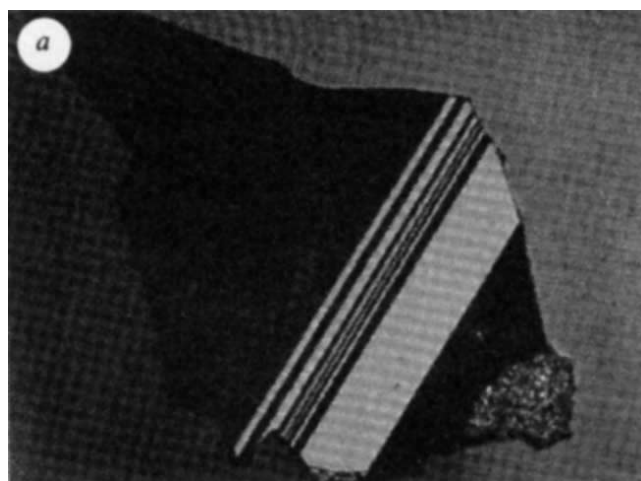


FIG. 2 a Polarized light micrograph of a partially twinned YBCO single crystal. The dark and bright areas correspond to different crystal orientations, and the lines between them are the twin boundaries. b, Fundamental Fourier component (in Gauss) of the response of the crystal shown in a to an applied a.c. field of frequency $\sim 1 \text{ Hz}$ and amplitude of 135 Oe , at a temperature $T = 46 \text{ K}$. Notice the preferential flux penetration along the twin boundaries. c, Phase lag (in radians) of the local field perturbation with respect to the applied oscillation, at $T = 56 \text{ K}$. The phase slope is inversely proportional to the perturbation propagation velocity, giving a ratio of 5:1 for the velocities along the twin boundary and in the bulk of the untwinned regions. The distance scale shown is in micrometres. The three-dimensional images in b and c are from the bottom of panel a, which is a two-dimensional view of the sample.

of 135 Oe and frequency 1.0 Hz, at a temperature of 46 K. This plot clearly shows that the magnetic flux leaks into the sample preferentially along the twin boundaries, which act like conduits for the flow of the flux. In Fig. 2c we show a plot of the phase lag in radians of the local response with respect to the applied oscillating field. The temperature for the lower set of data is 56 K. The phase slope is inversely proportional to the propagation velocity and suggests a ratio of 5:1 for the velocities along the twin boundary as compared to the untwinned regions. The data shown here are for a sample with a twin on an inside corner. The effect is enhanced by demagnetization effects for such a case, but it is seen for all twins, regardless of the sample shape. The detailed microscopic explanation is unclear at the moment, but it is at least plausible that the critical currents across a twin are reduced, allowing preferential flux penetration into the sample along a twin. Other explanations include strains near the twin or defects in the crystal itself which may have caused the twin to form there in the first place. The magnetic field is aligned to within 3° of the twins in the crystal. We have checked and this does not seem to be the cause of the preferential flow.

Figure 3 shows the results of a second kind of experiment: flux creep. Figure 3a is a polarized light micrograph of our sample, where the twins run at $\sim 45^\circ$ to the edge of the figure. The sample is zero-field cooled to a temperature of 44 K, where a field of 500 Oe is instantaneously applied and the response as a function of time is recorded. Figure 3b shows the field distribution 100 ms after applying the field; Figure 3c shows the response after 13 min. Again, the magnetic flux clearly leaks into the sample preferentially along the twin boundaries. In the bulk of

the sample (the red regions) even after 13 min very little of the field can be seen to have penetrated into the sample, whereas along the twins (the green spikes) the penetration into the sample can be seen after only 100 ms and is pronounced after 13 min. The direct measurement of the flux profile slopes gives a ratio of $\sim 3:1$ between untwinned and twinned regions, suggesting that the critical current across the twin boundaries is three times smaller than in the bulk of the material. This is confirmed by observations at low temperatures, where we have measured the thickness of the current layer necessary to achieve the full shielding of a small a.c. magnetic field, and found that penetration of the flux along the twins was a factor of four deeper than elsewhere.

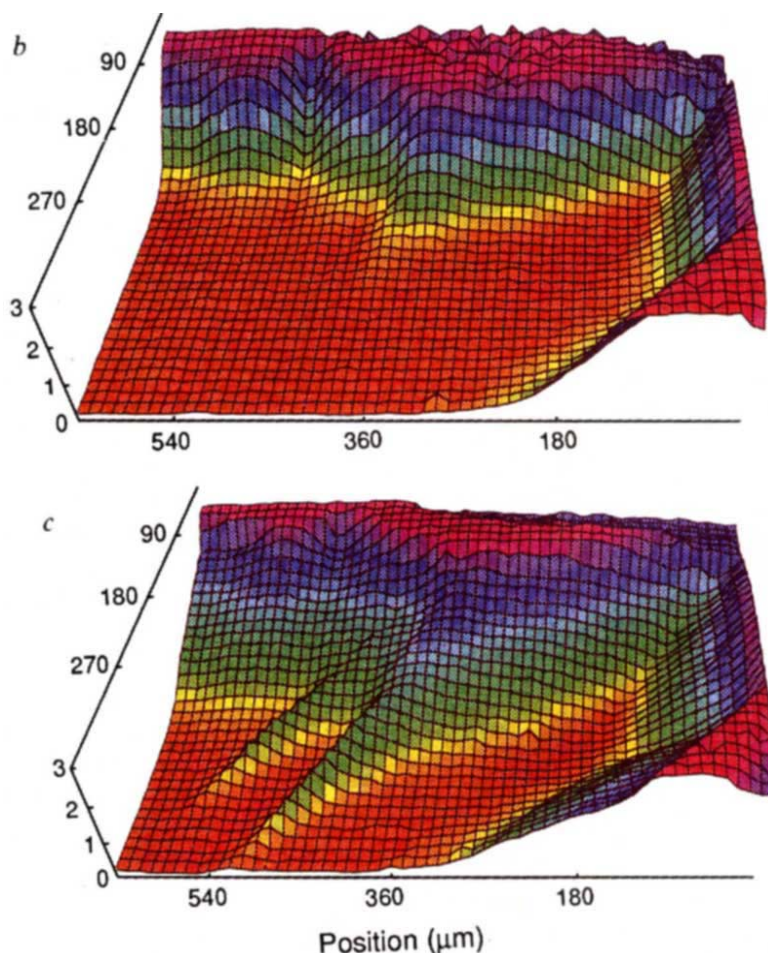
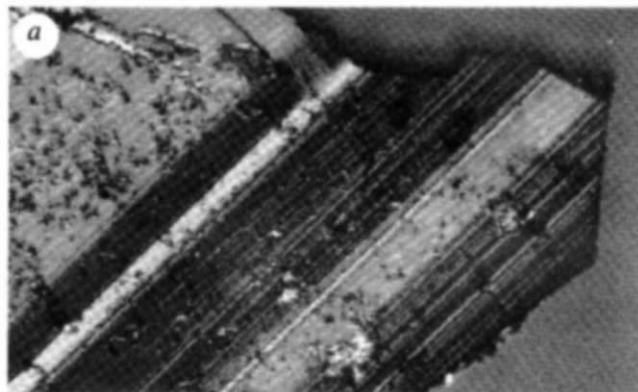


FIG. 3 Flux-creep measurements after zero-field cooling of the crystal shown in *a*, at $T = 44$ K and applied field of 500 Oe. *b* Shows the field distribution 100 ms after the field was turned on, and *c* shows the distribution after 13 minutes. The field scale is in arbitrary units and the distance scale is in micrometres. The three-dimensional images in *b* and *c* are seen as if from the bottom of *a*, which is a two-dimensional view of the sample.

We have also measured the flow of magnetization across a twin. It is known from magnetic decoration⁹, and measurements of transport¹⁰, magnetization¹¹ and torque¹², that in this direction the twins act as strong pinning sites, at least at high temperatures. Our experiments at lower temperatures confirm this. In cases where the magnetization in the bulk would like to flow perpendicular to the twin, we find that magnetization piles up at the twin. The twin boundary acts as a barrier to the flow.

Given that the twins act like conduits for the flow of magnetization into the interior of the sample but block the flow across the twins, our results suggest that bulk measurements are likely to be dominated by the twins both parallel to and perpendicular to the flow of the magnetization. This suggests the need for caution in interpreting such data as measuring the intrinsic magnetic response of the YBCO system. This observation is also relevant to the issue of critical currents in these materials. When a current flows in a superconductor, energy is dissipated only if the flux lines are able to move. How and where they pin is technologically important. We expect that spatially resolved techniques capable of monitoring the dynamics will help us to understand vortex liquids, solids and glasses, and how to pin them to obtain the highest critical currents¹³. □

Received 6 April; accepted 15 May 1992.

1. Bishop, D. J., Gammel, P. L., Huse, D. A. & Murray, C. A. *Science* **255**, 165–172 (1992).
2. Alers, P. B. *Phys. Rev.* **105**, 104–108 (1957).
3. DeSorbo, W. *Phys. Rev. Lett.* **4**, 406–408 (1960).
4. Indenbom, M. V. et al. *Physica C* **166**, 486–496 (1990).
5. Gotoh, S. et al. *Jap. J. appl. Phys.* **29**, L1083–L1085 (1990).
6. Raider, S. I. et al. *Appl. Phys. Lett.* **58**, 1676–1678 (1991).
7. Eschenfelder, A. H. *Magnetic Bubble Technology* (Springer, Berlin, 1981).
8. Rice, J. P. & Ginsberg, D. M. *J. Cryst. Growth* **109**, 432–435 (1991).
9. Dolan, G. J. et al. *Phys. Rev. Lett.* **62**, 827–830 (1989).
10. Kwok, W. K. et al. *Phys. Rev. Lett.* **64**, 966–969 (1990).
11. Swartzendruber, L. J. et al. *Phys. Rev. Lett.* **64**, 483–486 (1990).
12. Gyorgy, E. M. et al. *Appl. Phys. Lett.* **56**, 2465–2467 (1990).
13. Larbalestier, D. *Phys. Today* **44**, 74–82 (1991).

ACKNOWLEDGEMENTS. The work of J.P.R. and D.M.G. was supported by the NSF.

Flexible light-emitting diodes made from soluble conducting polymers

G. Gustafsson, Y. Cao, G. M. Treacy, F. Klavetter, N. Colaneri & A. J. Heeger

UNIAx Corporation, 5375 Overpass Road, Santa Barbara, California 93111, USA

THE recent fabrication of light-emitting diodes (LEDs) from conjugated polymers^{1,2} demonstrates the technological potential of this class of electronic materials. A variety of colours are possible, because the wavelength of luminescence emission can be chemically tuned during synthesis^{1–4}. In addition, the mechanical properties of polymers suggest that light-emitting structures can be made that are more flexible than their inorganic counterparts, provided appropriate materials can be found for the substrate and electrodes. Here we report the fabrication of a fully flexible LED using poly(ethylene terephthalate) as the substrate, soluble polyaniline as the hole-injecting electrode, a substituted poly(1,4-phenylene-vinylene) as the electroluminescent layer and calcium as the electron-injecting top contact. The structure is mechanically robust and may be sharply bent without failure. The LED is easily visible under room lighting and has an external quantum efficiency of about 1%. With a turn-on voltage for light emission of 2–3 V, the ‘plastic’ LED demonstrates that this unique combination of optical, electrical and mechanical properties can be used to make novel structures that are compatible with conventional devices.

An important advantage of the polymer LED over other electroluminescent devices is its ease of fabrication. Because the conjugated polymer active layer (or its precursor polymer) is soluble and can be deposited on a substrate from solution, simple processing techniques such as spin-casting or dipcoating can be used. The ability to apply the active luminescent polymer layer from solution makes it possible to fabricate large-area devices on flexible substrates.

In previously reported electroluminescent devices^{1–4} made from organic materials (with either organic molecules or conjugated polymers as the luminescent species), the light-emitting structures were not, in fact, mechanically flexible because the injecting contacts or the substrate (or both) used brittle materials. The most widely used transparent electrode material is indium/tin oxide, but as this is brittle, the electrode is easily shattered, losing both conductivity and optical clarity.

To avoid these problems, it would be desirable to use a conducting polymer which has been doped to the metallic state to yield a highly conducting and transparent thin film on a polymer substrate. Although transparent conducting polymer electrodes have been reported^{5–7}, they are polymerized *in situ*, in a matrix polymer or on the surface of a polymer substrate. Such processes are necessarily slow. Furthermore, they typically require a catalyst or an oxidizing agent, which affects the purity of the active semiconducting layer and the efficiency of the device.

Figure 1 shows schematically the structure of the flexible ‘plastic’ light-emitting diode. The device consists of a free-standing poly(ethylene terephthalate) (PET) film as substrate, a thin film of polyaniline (PANI) as the hole-injecting contact, a film of poly(2-methoxy, 5-(2'-ethyl-hexoxy)-1,4-phenylene-vinylene) (MEH-PPV) as the active electroluminescent layer, and a calcium electrode as the metallic electron-injecting contact.

Polyaniline films were spin-cast from solution in the conducting form (emeraldine salt) onto the PET substrate. Before use, the PET film (thickness 100 μm) was cleaned by boiling first in acetone and then isopropanol, and dried at 80 °C for 1 hour. The PANI solution was prepared using the method of ref. 8 which uses functionalized protonic acid solutes to dope polyaniline and render the resulting PANI complex soluble in common organic solvents. In such complexes⁸, the counter-ion, (M^+-R), contains the R functional group which is chosen to be compatible with nonpolar or weakly polar organic liquids.

An example of a transparent electrode for the flexible LED structure is camphor sulphonic acid (CSA) as the R functional group. After doping and complexation with CSA, the conducting PANI-complex is soluble in *m*-cresol and can be spin-cast onto

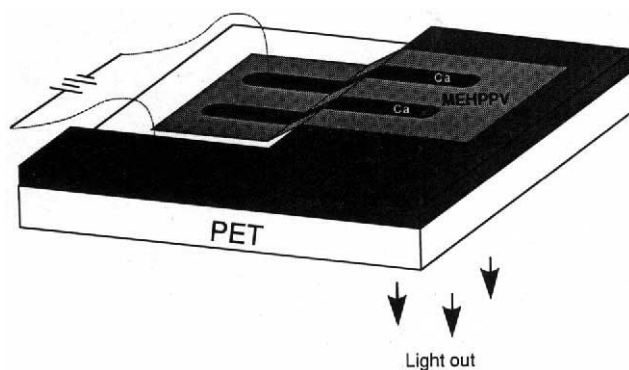


FIG. 1 Schematic diagram of the structure of a flexible light-emitting diode. The device is built on a PET substrate with successive layers as follows: PANI, then MEH-PPV, then Ca (see text). Film thicknesses are as follows: PANI-CSA, $\sim 0.5 \mu\text{m}$; MEH-PPV, 1,000–1,500 Å; calcium, $< 1,000 \text{ Å}$.

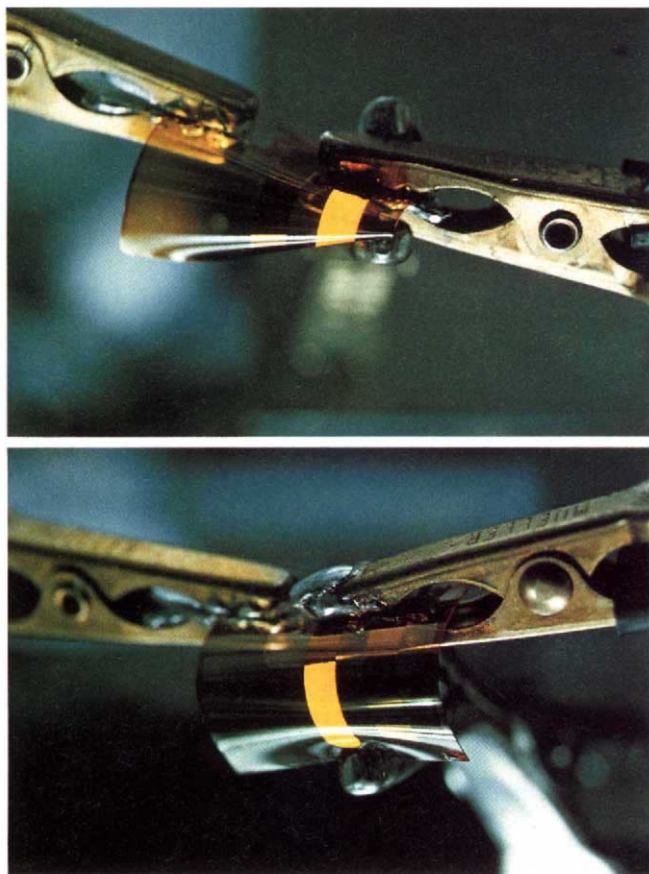


FIG. 2 Photographs of the flexible light-emitting diode at an applied bias voltage (positive on the PANI electrode) of 6 V. The structure of the device is shown in Fig. 1.

a variety of substrates, including glass and PET, to make optical-quality transparent films with low surface resistance⁹. Such films can achieve surface resistances lower than 100 Ω per square while retaining transmission greater than 70% at visible wavelengths greater than 475 nm (green to red)⁹.

The PANI-CSA films used for the hole-injecting contacts were spun from solution in meta-cresol (4% by weight) and dried at 60 °C for 12 hours. The thickness and accordingly the surface resistance of the PANI electrode are controlled by varying the spinning speed and/or the concentration of the PANI solution⁹; the films used in the flexible LEDs had a surface resistance of $\sim 200 \Omega$ per square.

Contact to the PANI electrode was made by a thin, vacuum-evaporated gold film. Although we used gold initially to make certain that the PANI electrode was not current-limiting, we find that excellent ohmic contacts are made to the transparent PANI electrode with silver paste or, simply, with an alligator clip. A layer of MEH-PPV¹⁰ was deposited on the PANI surface by spin-casting from solution of MEH-PPV in xylene (0.5 wt%). We saw no indication of dissolution of the PANI by the MEH-PPV solution. The two-component system was designed so that each component, the PANI-CSA complex and MEH-PPV, was separately soluble, but each was insoluble in the solvent that dissolved the other. The deposition of the MEH-PPV layer was carried out in nitrogen atmosphere in a controlled-atmosphere dry-box.

The calcium electron-injecting contact² was deposited by vacuum evaporation at a pressure below 10^{-6} torr. Because calcium oxidizes rapidly in air, the evaporator was mounted inside the dry box, and all physical measurements were carried out with the LEDs in the dry box. The electrode was initially passivated by over-coating with aluminium and then applying a protective polymer layer (with such protection, slow degradation occurs in air over a period of weeks).

Two flexible LEDs (under forward bias at 6 V) are shown in Fig. 2 with alligator clips used as contacts. The photographs are taken through the transparent PANI electrode; the calcium back-electrode serves as a mirror that reflects part of the emitted light. The light emitted by the device is visible under ordinary room lighting even at low bias voltages. As pure MEH-PPV is a high-resistivity semiconductor (bandgap ~ 2.1 eV), the emitting area is defined by the calcium contact. Patterning of the flexible devices is easily accomplished by masking during the evaporation. As the figure shows, the 'plastic' LED is flexible and can be curled and bent repeatedly (even folded back on itself in a sharp 180° bend) without failing or losing surface conductance, in contrast to indium/tin oxide-coated PET. The robust nature of the four-layer structure clearly proves the excellent mechanical adhesion of the PANI/MEH-PPV and PANI/PET contact (and, in addition, of the Ca/MED-PEV contact).

The PANI electrode has a light green colour, as it has a transmission window extending from ~ 475 nm. The absorption spectrum of the PANI electrode (Fig. 3) shows the characteristic features of the emeraldine salt form of PANI with a strong absorption peak at 440 nm and a broad absorption band extending into the infrared. The latter has been interpreted as resulting from the free-carrier absorption from the partially filled band in a polaronic metal¹¹. As can be seen from the spectrum in Fig. 3, the electroluminescence from the semiconducting MEH-PPV active layer falls right in the transmission window between the two principal absorptions of the PANI electrode.

Figure 4 shows the current-voltage ($I-V$) characteristic of a

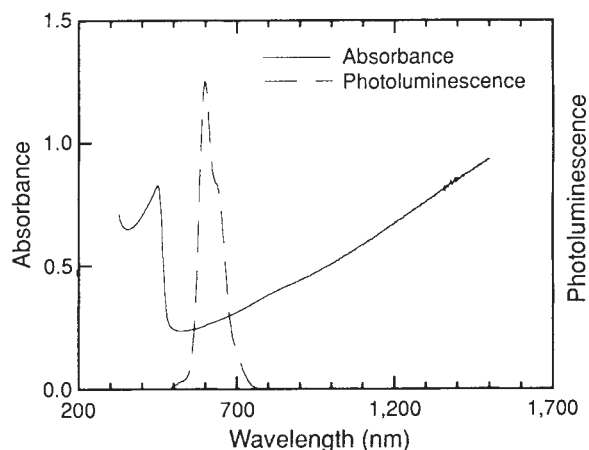


FIG. 3 The absorption spectrum of a PANI-CSA film compared with the photoluminescence spectrum of a MEH-PPV film.

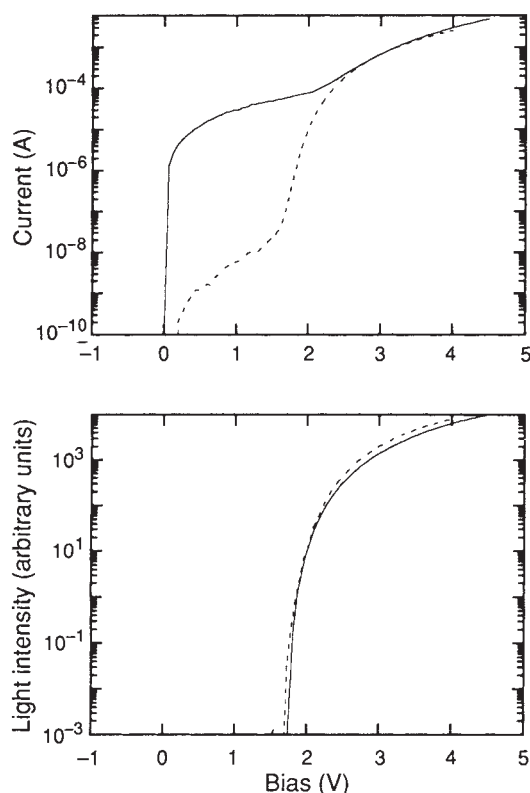


FIG. 4 The current (top) and electroluminescent intensity (bottom) as a function of applied bias voltage for light-emitting diodes with glass/ITO (---) and PET/PANI (—) hole-injecting electrodes. The two devices have equal areas, 0.1 cm².

flexible LED device, together with that of a device using indium/tin oxide (ITO) as the hole-injecting contact. At low bias (<1.8 V), the current is 10³–10⁴ times higher in the PANI/MEH-PPV/Ca device than in the ITO/MEH-PPV/Ca device. Above 1.8 V, however, in the range where the devices start to emit light, the currents through the two devices are similar. The luminescence intensities of the two devices (PANI or ITO electrode) are nearly identical; the quantum efficiencies (photons per electron) are ~1%.

Because the characteristics of the indium/tin oxide and the PANI devices are similar at high voltages, the two transparent electrode materials have similar hole-injecting properties. This can be rationalized by comparing their work functions. The work function of PANI as measured by photoelectron spectroscopy is 4.3 eV (P. Dannetun, personal communication), close to that of indium/tin oxide which has been reported¹² as 4.1 eV. We note, however, that the chemical stoichiometry and morphology of indium/tin oxide (and thus the work function) vary when such electrodes from different sources are compared. □

Received 14 February; accepted 16 April 1992.

- Burroughs, J. H. *et al.* *Nature* **347**, 539–541 (1990).
- Braun, D. & Heeger, A. J. *Appl. Phys. Lett.* **58**, 1982–1984 (1991).
- Burn, P. L. *et al.* *Nature* **356**, 47–49 (1992).
- Grem, G., Leditzky, G., Ulrich, B. & Leising, G. *Adv. Mat.* **4**, 36–38 (1992).
- Ojio, T. & Miyata, S. *Polym. J.* **18**, 95–98 (1986).
- Li, C. & Song, Z. *Synth. Met.* **40**, 23–28 (1991).
- Zhang, H. & Li, C. *Synth. Met.* **44**, 143–146 (1991).
- Cao, Y., Smith, P. & Heeger, A. J. *Synth. Met.* (in the press).
- Cao, Y., Treacy, G. M., Smith, P. & Heeger, A. J. *Appl. Phys. Lett.* (in the press).
- Wudl, F., Allemand, P. M., Srdanov, G., Ni, Z. & McBranch, D. in *Materials for Nonlinear Optics: Chemical Perspectives* (eds Marder, S. R., Sohn, J. E. & Stucky, G. D.) (683–686 American Chemical Society, Washington DC, 1991).
- Staflstrom, S. *et al.* *Phys. Rev. Lett.* **59**, 1464–1467 (1987).
- Feucht, D. L. *J. Vac. Sci. Tech.* **14**, 57–64 (1977).

Preparation and characterization of C₆₀Br₆ and C₆₀Br₈

Paul R. Birkett, Peter B. Hitchcock, Harold W. Kroto, Roger Taylor & David R. M. Walton

School of Chemistry and Molecular Sciences, University of Sussex, Brighton, BN1 9QJ, UK

BUCKMINSTERFULLERENE (C₆₀) is much more reactive than was originally anticipated, because resonance structures that place double bonds in the pentagonal rings are energetically unfavourable¹ and lead to restricted electron delocalization. C₆₀ therefore behaves as an electron-deficient 'super-alkene' rather than as a 'super-aromatic', as demonstrated by the addition of osmium tetroxide², platinum³, dipolar molecules⁴, alkyl radicals⁵ and amines⁶. Reaction of anions derived from C₆₀ with iodomethane results in up to 24 methyl groups becoming attached to the cage, although compounds containing either eight or six methyl groups are dominant in the mass spectrum⁷. Here we report the synthesis of crystalline C₆₀Br₆ (which forms magenta plates) and C₆₀Br₈ (dark brown prisms) by mixing solutions of C₆₀ with bromine. The structures of these compounds, which contain occluded bromine when precipitated from the brominating medium, have been determined by single-crystal X-ray diffraction. Reaction of C₆₀ with bromine in the absence of solvent gives C₆₀Br₂₄ (which also contains occluded bromine). We propose a configuration for C₆₀Br₂₄ (and sterically hindered C₆₀X₂₄ compounds) based on that for C₆₀Br₈. On being warmed in solution, C₆₀Br₆ rearranges and disproportionates to C₆₀Br₈, and all three bromo derivatives revert to C₆₀ on strong heating. These results provide an insight into the pattern of addition of bulky reagents to C₆₀.

Halogeno compounds are important synthetic intermediates, being reactive towards many nucleophiles⁸. Halogenation of

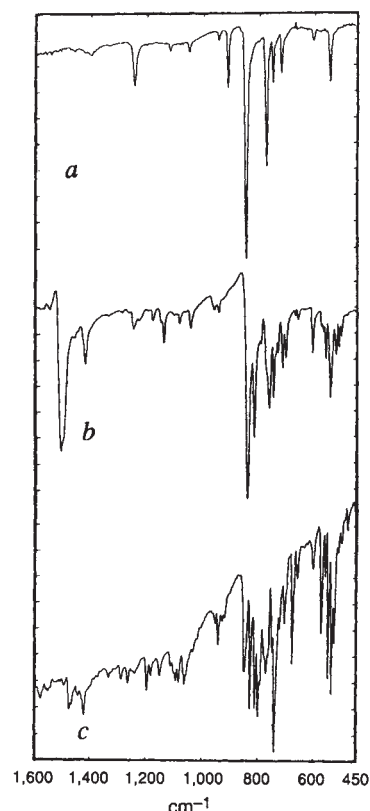


FIG. 1 Infrared spectra of (a) C₆₀Br₂₄; (b) C₆₀Br₈; (c) C₆₀Br₆. The strong C=S stretching band in spectrum b is due to occluded CS₂.

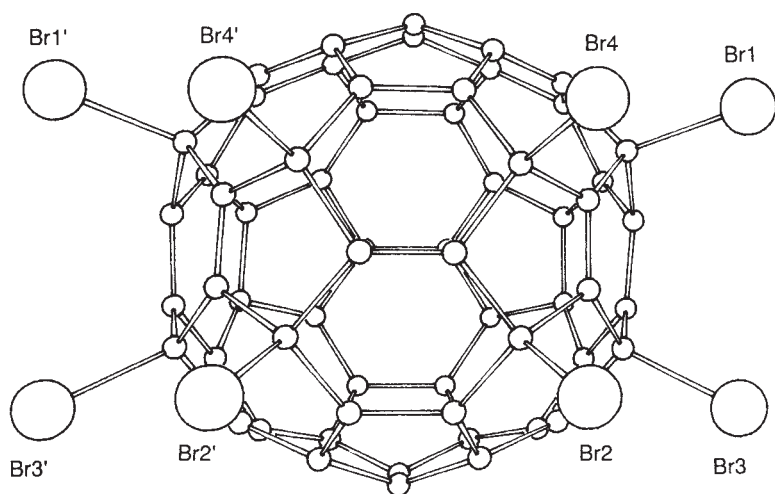


FIG. 2 Molecular structure of $C_{60}Br_8$. Crystals of the bromine-solvated molecule, $C_{60}Br_8 \cdot xBr_2$ (x is probably < 2) are orthorhombic, $Pnma$ (no. 62), $a = 24.482(8)$, $b = 16.895(4)$, $c = 11.343(3)$ Å, volume $U = 4,692$ Å³, $Z = 4$, $T = 173$ K, $R = 0.137$ for 803 reflections with $14 < \theta < 23^\circ$, and intensity $I > 3\sigma(I)$. The $C_{60}Br_8$ molecules lie on a crystallographic mirror plane. The Br_2 solvate molecules lie in channels parallel to the crystallographic c -axis and are only partially ordered.

C_{60} , for example fluorination^{9,10}, chlorination^{11,12} and bromination^{11,12}, has been achieved, but in an uncontrolled way and without structural characterization of the products. It is noteworthy that 24 chlorines¹¹ and 'about' 28 bromines¹² add to C_{60} . The fluoro compounds are susceptible to nucleophilic substitution^{13,14}, and replacement of the chloro group by methoxy has been reported¹¹. Mass spectrometric results provide evidence that the reaction of C_{60} with bromine and ferric chloride in benzene produces the species $C_{60}Ph_n$ ($n = 6, 8, 12$ and 16)¹⁵, indicating that Friedel-Crafts alkylation of benzene by brominated C_{60} intermediates had occurred. We therefore undertook a study of the bromination of C_{60} .

With neat bromine, C_{60} readily adds on 28 bromines (elemental analysis) to give a quantitative yield of yellow bromine-soluble microcrystals, which we believe to be $C_{60}Br_{24}$ containing two molecules of bromine (per fullerene) occluded in the lattice. When this compound is heated at $\sim 150^\circ\text{C}$ during 24 h all the bromine is eventually lost (see also ref. 11) in what seems to be a cascade process, because infrared monitoring indicates the presence, at any given time, of a mixture of C_{60} and the fully brominated starting material. Pure C_{70} takes up a similar number of bromine atoms and likewise loses them on heating. $C_{60}Br_{24}$ has a simple infrared spectrum (Fig. 1a) indicating a highly symmetrical structure, as proposed below. The main bands occur

at 546, 606, 720, 751, 776, 849, 912, 946, 1,050, 1,117, 1,244 and $1,400\text{ cm}^{-1}$.

On brominating C_{60} in CS_2 , dark brown crystals were deposited from solution in 80% yield during 24 h. (No reaction occurs at low bromine concentration, indicating a high-order process). The crystals were separated, and their structure determined by single-crystal X-ray diffraction using data collected on an Enraf-Nonius CAD4 diffractometer with Mo $K\alpha$ radiation. This showed the crystals to be $C_{60}Br_8$, having two identifiable molecules of occluded bromine (and other very disordered bromine or CS_2 molecules) per fullerene molecule in the lattice (Figs 2 and 3). Because of the contribution to the scattering of the extensively disordered component, there were problems with the refinement. Ultimately, we only used data with $\sin \theta/\lambda > 0.35$, which are less affected, and we placed constraints on the geometry of the $C_{60}Br_8$ molecule. The sets of bonds that were constrained to be equal and the values to which they refined are: C-Br, $1.97(5)$ Å; C-C(Br), $1.52(3)$ Å; other C-C bonds within pentagons, $1.44(3)$ Å; other C-C bonds between pentagons, $1.40(5)$ Å. The two isolated double bonds were not constrained and have values of $1.27(15)$ Å and $1.30(15)$ Å.

$C_{60}Br_8$ was also obtained in 58% yield by brominating in chloroform. It has low solubility in most organic solvents but is soluble in bromine. The infrared spectrum of material obtained

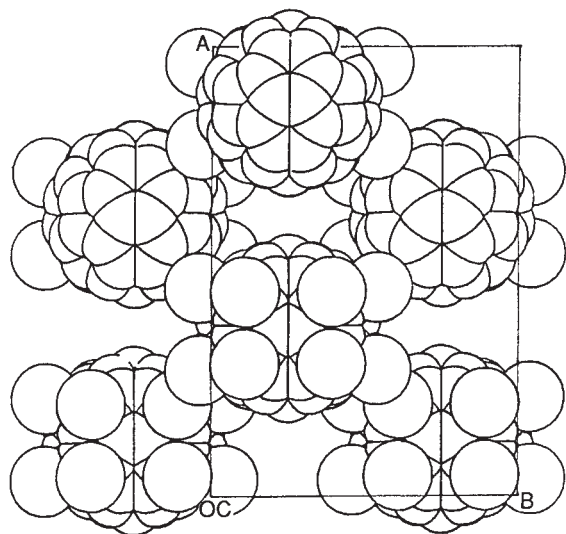


FIG. 3 Crystal structure of $C_{60}Br_8$ showing the channels, parallel to the c -axis, into which bromine molecules fit.

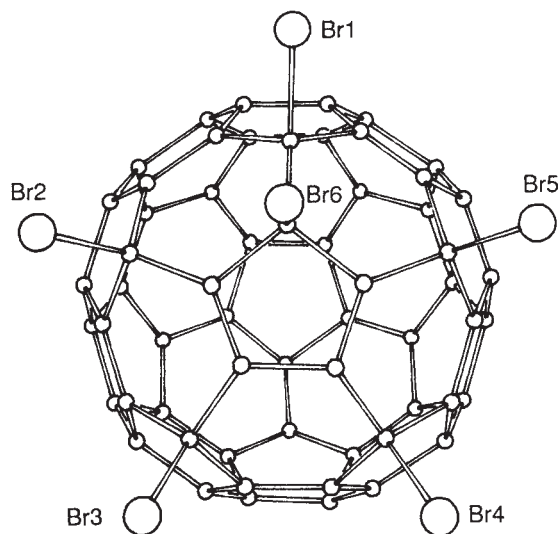


FIG. 4 Molecular structure of $C_{60}Br_6$. Crystals of the bromine-solvated molecule, $C_{60}Br_6 \cdot xBr_2$ ($x \approx 1$) are monoclinic, $P2_1/c$ (no. 14), $a = 18.190$ (3), $b = 11.898$ (5), $c = 18.117$ (5) Å, $\beta = 96.60$ (2)°, $U = 3.895$ Å³, $Z = 4$, $T = 295$ K, $R = 0.092$ for 2,328 reflections with $2 < \theta < 25^\circ$ and $I > 3\sigma(I)$. Br_2 solvate molecules lie on inversion centres whereas the $C_{60}Br_6$ molecule is in a general position.

using CS_2 as solvent is shown in Fig. 1b, and is more complex than that of $C_{60}Br_{24}$ because of the lower symmetry. The main bands occur at 516, 526, 546, 563, 610, 706, 718, 750, 766, 820, 845, 947, 963, 1,047, 1,086, 1,142, 1,182, 1,250 and $1,422\text{ cm}^{-1}$.

Bromination in either benzene or carbon tetrachloride produces (in 54 and 92% yields, respectively) magenta plates. The structure of a sample obtained by bromination in benzene was shown by X-ray diffraction to be $C_{60}Br_6$, containing one molecule of bromine (per fullerene molecule) occluded in the lattice (Fig. 4). The C-Br6 bond (2.03 (2) Å) is longer than the other C-Br bonds, which average 1.96 (3) Å. The butadiene fragment of the central pentagon shows individual bond lengths of 1.31 (4), 1.47 (3) and 1.36 (3) Å. C-C(Br) single bonds average 1.53 (3) Å; the remaining C-C bonds in pentagons

average 1.45 (3) Å and the remaining interpentagonal bonds average 1.38 (3) Å. The average distance to the centre of the cluster is 3.52 (4) Å for sp^2 -hybridized carbons (compare this with 3.512 Å for the osmium tetroxide derivative¹⁶ and 3.80 (4) Å for the sp^3 -hybridized carbons that bear bromine atoms).

$C_{60}Br_6$ is moderately soluble in organic solvents. Of the three bromo derivatives, it has the lowest symmetry and consequently the most complex infrared spectrum (Fig. 1c). The main bands occur at 458, 484, 529, 538, 551, 562, 575, 606, 661, 679, 708, 742, 751, 801, 812, 829, 850, 945, 1,065, 1,085, 1,152, 1,198, 1,265, 1,291, 1,421 and $1,473\text{ cm}^{-1}$.

Each of the bromo derivatives loses bromine on heating in the absence of solvent ($C_{60}Br_{24}$ is the more stable), so mass spectra cannot be obtained. When $C_{60}Br_6$ is heated in either carbon tetrachloride or benzene, it disproportionates to give $C_{60}Br_8$ and C_{60} . Formation of $C_{60}Br_8$ may involve a sequence of three 1,3-allylic bromine shifts initiated by the less-securely attached central bromine atom (greater C-Br bond length), followed by bromination with displaced or occluded bromine. The instability of $C_{60}Br_6$ and $C_{60}Br_8$ may reflect the presence of four and two double bonds in pentagonal rings in their respective structures; moreover $C_{60}Br_6$ contains an eclipsing steric interaction. The higher stability of $C_{60}Br_{24}$ is reasonable given the considerable bond reorganization required to regain the C_{60} structure, leading to a higher-energy transition state. The isolation of different derivatives in different solvents may arise from selective precipitation of the insoluble bromo derivatives, which then inhibits further reaction.

We note the propensity for either six or eight groups to add in bromination, phenylation and methylation. A different addition pattern may be involved in reaction with less-bulky reagents¹⁷. The formation of derivatives containing either 12 or 16 phenyl groups may involve repetition of the 6- and 8-fold addition patterns, and if so, many possibilities can be envisaged in each case. Addition of three batches of eight groups in the eightfold pattern described can, however, lead only to $C_{60}X_{24}$ (where X is Br or other bulky group), the structure of which is shown in the Schlegel diagram, Fig. 5.

Note on nomenclature. Our results highlight the need to have a nomenclature for describing fullerene derivatives. One possibility is to label in a spiral from one pentagonal cap through to its antipode. According to this scheme, the derivatives we have isolated are then 1,6,9,12,15,18-hexabromo-, and 1,3,6,11,13,18,28,31-octabromofullerene-60. The bromo compound corresponding to the structure in Fig. 5 is thus 1,3,6,8,11,13,16,18,21,23,26,28,31,33,36,38,41,44,46,49,51,54,57,60-tetracosabromofullerene-60.

Note added in proof: Since this work was completed, Tebbe *et al.*¹⁸ have published a single-crystal X-ray structure of $C_{60}Br_{24}$.

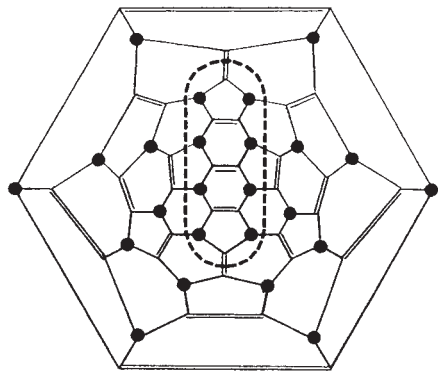


FIG. 5 Schlegel diagram for the proposed structure of $C_{60}X_{24}$, where X is bulky and located at the sites having filled circles. One of the sets of eightfold additions is shown by the dotted line.

Received 18 March; accepted 1 May 1992.

1. Taylor, R. *Tetrahedron Lett.* 3731-3734 (1991); *J. chem. Soc. Perkin Trans. 2*, 3-4 (1992).
2. Hawkins, J. M. *et al. J. org. Chem.* **55**, 6250 (1990).
3. Fagan, P. J., Calabrese, J. C. & Malone, B. *Science* **255**, 1160 (1991); *J. Am. chem. Soc.* **113**, 9408-9409 (1991).
4. Suzuki, T., Li, Q., Khemani, K. C., Wudl, F. & Almarsson, Ö. *Science* **254**, 1186-1188 (1991).
5. Krusic, P. J., Wasserman, E., Keizer, P. N., Morton, J. R. & Preston, K. F. *Science* **254**, 1183-1185 (1991).
6. Hirsch, A., Li, Q. & Wudl, F. *Angew. Chem. Int. Edn Engl.* **30**, 1309-1310 (1991).
7. Bausch, J. W. *et al. J. Am. chem. Soc.* **113**, 3205 (1991).
8. de la Mare, P. B. D. & Swedlund, P. E. in *Chemistry of the Functional Groups*, Vol. 16, Ch. 7 (ed. Patai, S.) (Wiley, Chichester, 1973).
9. Holloway, J. H. *et al. J. chem. Soc. Chem. Commun.*, 966-969 (1991).
10. Selig, H. *et al. J. Am. chem. Soc.* **113**, 5475 (1991).
11. Olah, G. A. *et al. J. Am. chem. Soc.* **113**, 9385-9387 (1991).
12. Tebbe, F. N. *et al. J. Am. chem. Soc.* **113**, 9900-9901 (1991).
13. Taylor, R. *et al. Nature* **355**, 27 (1991).
14. Taylor, R. *et al. J. chem. Soc. Chem. Commun.*, 665-667 (1992).
15. Taylor, R. *et al. J. chem. Soc. Chem. Commun.*, 667-668 (1992).
16. Hawkins, J. M. *et al. Science* **252**, 312 (1991).
17. Taylor, R. *J. chem. Soc. Perkin Trans. 2*, 453-455 (1992).
18. Tebbe, F. N. *et al. Science* **256**, 822-825 (1992).

ACKNOWLEDGEMENTS. We thank BP Ltd, ICI Ltd and the Royal Society for financial support.

Eddy momentum flux and its contribution to the Southern Ocean momentum balance

Rosemary Morrow*, John Church†‡, Richard Coleman§, Dudley Chelton|| & Neil White†

* Marine Studies Centre, University of Sydney, New South Wales 2006, Australia

† Division of Oceanography, CSIRO Marine Laboratories, Hobart, Tasmania 7001, Australia

§ School of Civil and Mining Engineering, University of Sydney, New South Wales 2006, Australia

|| College of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

‡ Also at: Co-operative Research Centre for the Antarctic and Southern Ocean Environment, University of Tasmania, Tasmania 7001, Australia

A LARGE amount of momentum is transferred to the Southern Ocean by strong westerly winds. Analytical and numerical models have suggested that transient eddies may be important in transporting this momentum away from the region of wind forcing, either horizontally¹ or vertically downwards where it is balanced by bottom topographic drag^{2–5}. There are, however, few long-term *in situ* observations of horizontal eddy momentum flux^{6–8}, and no large-scale measurements of vertical eddy fluxes, to test these models. As a result, the momentum balance of the Antarctic circumpolar current (ACC) remains uncertain, and the role of eddies controversial. Here we use Geosat satellite altimeter data to resolve directional eddy kinetic energy and horizontal eddy momentum flux in the ACC on fine spatial and temporal scales. The complex spatial distribution of surface eddy momentum flux is strongly influenced by bottom topography. The horizontal eddy momentum flux tends generally to concentrate the mean flow, although some regions of divergence are observed. Our results show that the zonally averaged horizontal eddy momentum flux from transient eddies is an order of magnitude too small, and in the wrong direction to directly balance the eastward momentum input from wind¹.

We use two years of Geosat altimeter data to calculate the variance in surface velocity at crossover points, where time series of the two components of velocity from ascending and descending groundtracks are available. The absolute velocity cannot be determined accurately from altimetry because of uncertainty in the geoid⁹. The time-variable (eddy) component of cross-track geostrophic velocity is, however, easily obtained, after standard corrections have been applied, and the mean height and orbit error removed¹⁰. The two velocity components at a crossover

point are not measured concurrently, because of the sampling characteristics of the satellite orbit, and may be separated in time by up to 8.5 days. To minimize this 'sampling error' we interpolate the velocities from ascending and descending groundtracks to the midpoint in time. We then use a simple geometric transformation to obtain the velocity in orthogonal north and east components (v' , u')¹¹. At 65° S, the angle between the ascending and descending groundtracks is almost orthogonal. As the ground track angle (relative to north) decreases towards the Equator, the zonal flow becomes better resolved, and the meridional flow becomes poorly resolved. For this reason we have confined our study to the Southern Ocean region south of 30° S, where both components are well resolved. Although the altimeter velocities are strictly surface values, previous studies have noted the high vertical coherence of currents in the Southern Ocean^{7,12} and stressed the importance of the barotropic and first baroclinic mode¹³. Thus our results are likely to be representative of the upper layers.

Figure 1 shows velocity-variance ellipses calculated from ensemble-averaged statistics over 3° × 3° bins. The large-scale structure is consistent with drifter studies at a coarser 5° × 5° resolution^{14,15}, although the regular altimeter sampling provides more degrees of freedom over the same time period. In the analysis, each crossover point is treated independently, and at each crossover there is one degree of freedom every 17 days (the repeat period of Geosat). The variance ellipses are mostly isotropic at this resolution, except in the high eddy energy regions of the Agulhas retroflection, the East Australian current and the Brazil current, where meanders in the mean current affect the orientation of the variance ellipses. Anisotropic ellipses are also observed where the mean current interacts with bottom topography, for example across the Madagascar Ridge at 40° E and the Macquarie Ridge at 160° E. There is generally little eddy activity over topographic ridges and shallower plateaus, whereas the higher eddy energy occurs in the deep basins, next to large topographic features. These results are consistent with numerical models^{4,16} and earlier altimetric studies based on variability in the sea surface height^{9,17}.

The advantage of the present technique is the ability to resolve the small horizontal surface features (of the order of 100 km), over an ocean basin, which cannot be achieved with mapped altimeter data on larger scales^{9,18}. The fine resolution is especially important in the Southern Ocean where abrupt changes in topography cause a similar response in the eddy field. Figure 2a shows an enlargement of the region southeast of Australia, with variance ellipses derived from the individual Geosat crossover points. At this resolution, the anisotropic nature of the eddy variance is more apparent. In regions of high eddy energy surrounding the Macquarie Ridge, the variance ellipse orientation is strongly steered by topography, especially along the

TABLE 1 Eddy flux statistics

Latitude (° S)	Longitude (° E)	Depth (m)	Eddy kinetic energy			Eddy momentum flux		
			Current meter (cm ² s ⁻²)	Altimeter (cm ² s ⁻²)	Ratio	Current meter (cm ² s ⁻²)	Altimeter (cm ² s ⁻²)	Ratio
Agulhas Array ⁸								
38.0	18.5	195	1,346	2,124	1.58	31.6	71	2.25
37.9	21.1	210	936	1,349	1.44	87.5	769*	8.79
37.2	23.0	194	407	679	1.67	82.8	37	0.45
35.9	27.0	210	464	999	2.15	104.9*	41	0.39
35.0	26.0	203	316	601	1.90	108.1*	83*	0.77
Ridge Array ⁷								
49.4	189.5	1,055	169	228	1.35	52.6	56*	1.06

Comparison of eddy flux statistics from surface measurements of Geosat altimeter data (November 1986–November 1988) with available *in situ* current-meter data at two Southern Ocean locations. The Agulhas Array was deployed for two years from March 1985, and the Ridge Array, south-east of New Zealand, for 17 months from November 1978. The surface altimeter results were bilinearly interpolated to the current-meter locations from the four surrounding crossover locations.

* These eddy momentum flux values are statistically significant at the 95% level⁷.

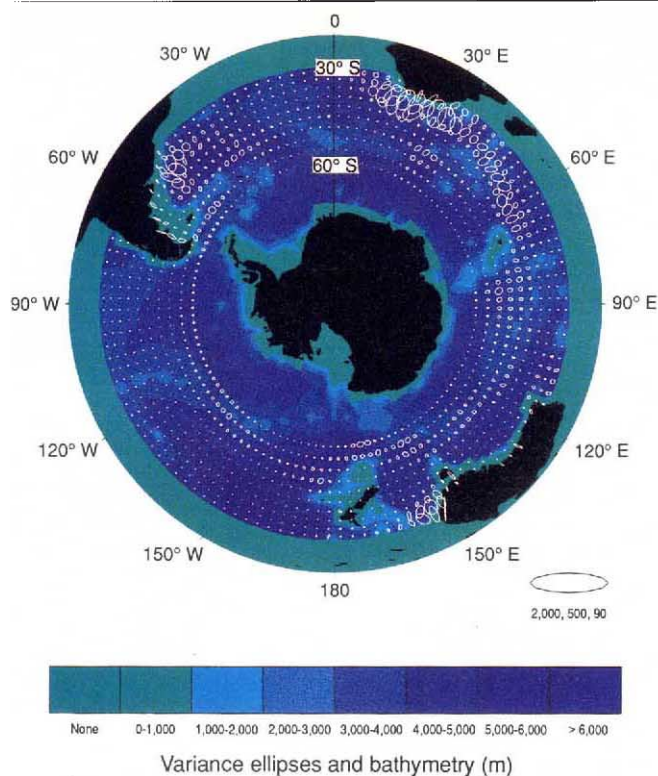


FIG. 1 Velocity variance ellipses on a $3^\circ \times 3^\circ$ grid for the Southern Ocean from 30°S to 60°S . Colour contours denote bottom bathymetry: regions with water depth cm^2s^{-2} less than 3,000 m are shaded light blue/green. The scale ellipse has a major axis of $2,000\text{ cm}^2\text{s}^{-2}$, minor axis of $500\text{ cm}^2\text{s}^{-2}$ and is oriented at 90° from north. Ellipses are calculated from the time series of u' , v' components and give a simple description of the direction and magnitude of the principal axes of velocity variance. The orientation of the ellipses is defined such that orthogonal velocity components parallel to the axes of the ellipse are uncorrelated over the duration of the Geosat data set. Ellipses oriented towards the northeast (southeast) have a positive (negative) eddy momentum flux, $\overline{u'v'}$. Each ellipse has on average 190 degrees of freedom.

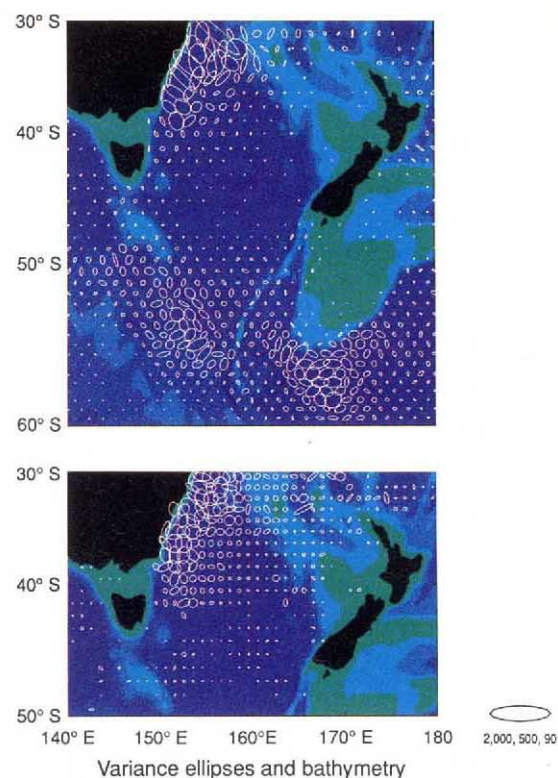


FIG. 2 Comparison of velocity variance ellipses between Geosat data and surface drifter data, for the Tasman Sea region southeast of Australia. *a*, The Geosat velocity variance calculated from the two-year time series of u' , v' components at crossover points, from November 1986. There are a maximum of 43 degrees of freedom at each crossover, assuming independent measurements every 17 days. The average number of degrees of freedom is 32, because of missing data. Only velocity variance ellipses above the estimated noise level of the altimeter are shown. *b*, Surface drifter variance ellipses from 30 years of data ensemble averaged on a 1° grid. The scale ellipse and the colour contours are described in Fig. 1.

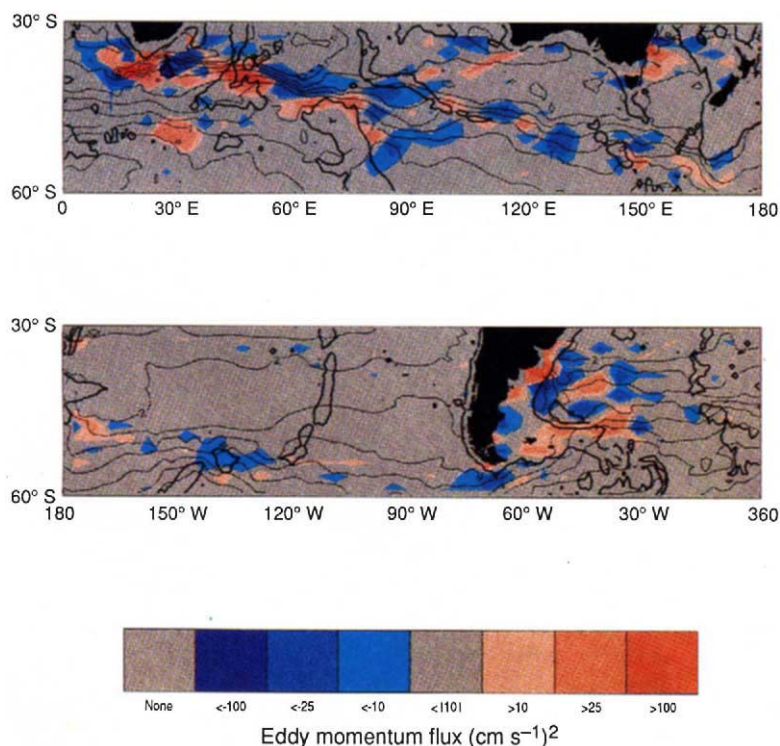


FIG. 3 Eddy momentum flux (or surface horizontal Reynolds stresses) in cm^2s^{-2} for the Southern Ocean from 30°S to 60°S on a 2° latitude by 5° longitude grid. Regions where there are positive (negative) fluxes of momentum are shaded red (blue). We estimate that each $2^\circ \times 5^\circ$ box has on average 218 degrees of freedom. As a result, most $\overline{u'v'}$ values less than $25\text{ cm}^2\text{s}^{-2}$ are below the 95% significance level⁷, although some regions of lower magnitude, such as the area of $\overline{u'v'}$ convergence downstream of the Pacific Mid-Ocean Ridge at 100°W , are significant at this level. The thick black line denotes the 3,000-m depth contour. The thin black lines show mean dynamic topography²⁴ with contour intervals of 20 cm. Close contours indicate the mean position and strength of the ACC.

southeast escarpment of Campbell Plateau. Neighbouring ellipses appear coherent along the axis of the ACC, but show distinct orientation changes across the ACC.

The variance ellipses in the East Australian current region (Fig. 2a) agree well with surface drifter data, shown on a $1^\circ \times 1^\circ$ grid in Fig. 2b (G.C. and J.P., manuscript in preparation). Near the coast, both data sets have elongated ellipses oriented parallel to the coast, which turn to the southeast and become more isotropic farther offshore. The eastward extent of the high eddy activity is limited by the Lord Howe rise at 160° E. The East Australian current turns east from the coast around 35° S and forms a series of north-south meanders between 152 and 164° E^{19,20} which are reflected in the orientation of the variance ellipses. Sixty per cent of the drifter measurements are in the latitude band from 30° to 40° S, and here the mean eddy kinetic energy is $341 \text{ cm}^2 \text{ s}^{-2}$ for the drifters and $365 \text{ cm}^2 \text{ s}^{-2}$ for the altimeter results. The spatial distribution of the velocity variance shows good agreement, given the very different measurement techniques, and the fact that the data sets span different periods and have different spatial coverage.

Measurements of surface eddy momentum flux ($\overline{u'v'}$) are important for understanding how eddies may act to transport momentum into or away from the mean flow, and the role of topography in this interaction. The interaction of the mean current with topography in the Southern Ocean produces a complex pattern in the $\overline{u'v'}$ field (Fig. 3). Table 1 shows a comparison of long-term current meter records with the Geosat eddy flux statistics, interpolated onto the current meter locations. The current meter records are measured at depths of 200 and 1,000 m, so the surface amplification of the Geosat velocity variance is expected. The eddy kinetic energy field tends to be spatially coherent, and the ratio of altimeter to current-meter results remains fairly constant. In contrast, the $\overline{u'v'}$ field has less spatial coherence between crossover points, especially in the complex Agulhas region, resulting in scatter when comparing interpolated altimeter results to the current-meter measurements. Upstream from the Agulhas Plateau at 27° E, all the current arrays show a northward flux of eddy momentum of $\sim 100 \text{ cm}^2 \text{ s}^{-2}$ at the shallowest depth of 200 m (ref. 8), which corresponds to the large positive maxima in Fig. 3. There is similar agreement for the positive maxima southeast of New Zealand⁷. Given the complex spatial distribution of $\overline{u'v'}$, the two data sets show reasonable agreement.

The topographic influence on the eddy momentum flux field is clearly seen in the Agulhas region south of Africa. There is a northward flux of eastward momentum where the mean flow is concentrated to the north of ridges such as the Del Cano Rise at 40° E. In the deep basins downstream the mean current is deflected southwards⁴ and there is an accompanying southward flux of eastward momentum. There are large areas where there is a southward flux of momentum to the north of the mean flow, and a northward flux of momentum south of the mean flow. This is particularly evident in the Agulhas region from 30° E to $\sim 80^\circ$ E, and near the Pacific Mid-Ocean Ridge at 140° W. This convergence of eddy momentum flux has been predicted in eddy-resolving models of the ACC³ and also observed in the Kuroshio Current^{18,21} and the Gulf Stream²². In numerical models, transient eddies generated by baroclinic instabilities of the mean flow frequently act to converge eddy momentum flux and intensify the mean jet, as observed in the Geosat results. Figure 3 also shows isolated regions of $\overline{u'v'}$ divergence, for example, downstream of the Pacific Mid-Ocean Ridge around 110° W.

A calculation by Gill¹ suggests that to balance an eastward wind stress of 1 dyn cm^{-2} over the circumpolar region requires a net meridional eddy flux of zonal momentum of $100 \text{ cm}^2 \text{ s}^{-2}$, directed away from the circumpolar band. Our results show a zonally averaged divergent eddy momentum flux of $\sim 5 \text{ cm}^2 \text{ s}^{-2}$ out of the circumpolar band. But the ACC is not strictly zonal, and a calculation along streamlines parallel to the axis of the

ACC gives a convergent flux of eddy momentum of $\sim 10 \text{ dyn cm}^{-2}$. The difference in magnitude between Gill's theory and our observations may be even larger, as the wind stress forcing is closer to 2 dyn cm^{-2} in this band²³, and our surface measurements are likely to overestimate the depth-averaged eddy momentum flux.

This analysis of altimeter data demonstrates that global-coverage satellite data can be used to examine important physical processes, especially in the vast Southern Ocean where shipboard observations are sparse and difficult to acquire. Although we have shown that the transient eddies serve to accelerate the mean flow of the ACC rather than balance the momentum input by the wind, eddies (both transient and standing) may still play an important part in the momentum balance by transferring zonal momentum downwards, to be balanced by topographic form drag^{3-5,16}. □

Received 16 December 1991; accepted 14 April 1992.

- Gill, A. E. *J. Fluid Mech.* **32**, 465–488 (1968).
- Munk, W. H. & Palmén, E. *Tellus* **3**, 53–55 (1951).
- McWilliams, J. C., Holland, W. R. & Chow, J. H. *S. Dynam. Atmos. Oceans* **2**, 213–291 (1978).
- Wolff, J.-O., Maier-Reimer, E. & Olbers, D. J. *J. phys. Oceanogr.* **21**, 236–264 (1991).
- Johnson, G. C. & Bryden, H. L. *Deep Sea Res.* **36**, 39–53 (1989).
- Piola, A. R., Figueroa, H. A. & Bianchi, A. A. *J. geophys. Res.* **92**, 5101–5114 (1987).
- Bryden, H. L. & Heath, R. A. *Prog. Oceanogr.* **14**, 65–87 (1985).
- Luyten, J. et al. *Woods Hole Oceanogr. Inst. Tech. Rep.* WHOI-90-30 (1990).
- Chelton, D. B., Schlax, M. G., Witter, D. L. & Richman, J. G. *J. geophys. Res.* **95**, 17877–17903 (1990).
- Van Gysen, H., Coleman, R., Morrow, R. A., Hirsch, B. & Rizos, C. *J. geophys. Res.* **97**, 2265–2277 (1992).
- Parke, M. E., Stewart, R. H., Farless, D. L. & Cartwright, D. E. *J. geophys. Res.* **92**, 11693–11707 (1987).
- Sciremammano, F. Jr. *J. phys. Oceanogr.* **9**, 221–226 (1979).
- Inoue, M. *J. phys. Oceanogr.* **15**, 1157–1181 (1985).
- Daniault, N. & Menard, Y. *J. geophys. Res.* **90**, 11877–11899 (1985).
- Patterson, S. L. *J. phys. Oceanogr.* **15**, 865–884 (1985).
- Treguier, A. M. & McWilliams, J. C. *J. phys. Oceanogr.* **20**, 321–343 (1990).
- Sandwell, D. T. & Zhang, B. *J. geophys. Res.* **94**, 17971–17984 (1989).
- Tai, C.-K. & White, W. B. *J. phys. Oceanogr.* **20**, 1761–1777 (1990).
- Coleman, R. *Austr. J. Mar. Freshwater Res.* **35**, 619–633 (1984).
- Mulhearn, P. J. *J. phys. Oceanogr.* **17**, 1148–1155 (1987).
- Nishida, H. & White, W. B. *J. phys. Oceanogr.* **12**, 160–170 (1982).
- Schmitz, W. J. *J. phys. Oceanogr.* **12**, 208–210 (1982).
- Trenberth, K. E., Large, W. G. & Olsen, J. G. *J. phys. Oceanogr.* **20**, 1742–1760 (1990).
- Gordon, A. L. & Molinelli, E. J. *Southern Ocean Atlas: Thermohaline and Chemical Distributions and the Atlas Data Set* (Columbia Univ. Press, New York, 1982).

ACKNOWLEDGEMENTS. We acknowledge M. Schlax's contribution to analysis of the altimeter data, and thank G. Cresswell and J. Peterson for the surface drifter data. We thank B. Douglas and his group for support and assistance. The work was partly supported by a grant from the Australian Research Council and the Institute of Natural Resources and Environment, CSIRO Climate Change Programme.

Global propagation of interannual fluctuations in atmospheric angular momentum

J. O. Dickey*, S. L. Marcus* & R. Hide†

* Space Geodetic Science and Applications Group, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA

† Meteorological Office Unit, Robert Hooke Institute, The Observatory, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK

THE El Niño Southern Oscillation (ENSO), a climate fluctuation that recurs on a 2–7-yr timescale, is associated with persistent large-scale fluctuations in the dynamical behaviour of the global atmosphere–ocean system¹. Here we present a study of the latitudinal redistribution of angular momentum within the atmosphere from 1976 to 1991. We observe slow, global-scale coherent poleward propagation of atmospheric angular-momentum fluctuations on interannual timescales. These originate in equatorial regions, where they lead the main atmospheric anomalies of the ENSO cycle by nearly two years; they penetrate to latitudes higher than 60° in both hemispheres, where they lag behind the ENSO cycle by about four years. We can also distinguish the bimodality

of the ENSO phenomenon, with a low-frequency component centred at a period close to 4.2 years and a high-frequency component centred near 2.4 years. Each of the two components has a distinct latitudinal propagation pattern. In the period studied, strong El Niño and related La Niña climatic events occur when these components add constructively.

The axial angular momentum of the atmosphere is a useful diagnostic measure of the state of the general circulation²; both the ENSO cycle and the intraseasonal oscillation (40–50 days), for example, have clear spectral signatures in this scalar quantity^{3–7}. Daily atmospheric angular momentum (AAM) determinations are available from 1976 onwards from the National Meteorological Center operational analysis. Before 1976, fluctuations in length of day may be used as a proxy for AAM variations⁷. Previous studies have indicated a strong correlation between interannual variations in the Southern Oscillation Index and length of day^{8–11}, and by implication the global AAM. ENSO events are associated with positive length-of-day anomalies, implying a lower rotation rate of the solid Earth and, by conservation of angular momentum, a positive AAM anomaly.

Here we use the longest AAM time series now available, from the National Meteorological Center, which covers 15.5 yr (January 1976–June 1991) and is based on zonal wind data from the surface up to the 100 mbar pressure level¹². The atmosphere is divided into 46 equal-area latitude belts, extending from -90° to $+90^\circ$, with three-day averages considered. Figure 1a shows a contour plot of the spectral power in the AAM data set for periods from one to ten years, formed by combining periodograms for the 46 individual belts in the latitude-frequency domain. Strong maxima are found in the subtropics of both hemispheres at periods of 3–5 yr, a typical recurrence time for moderate El Niño events¹. In the tropics, lower power levels on these timescales may be due to a baroclinic structure for low-frequency wind variations^{13,14}, although larger AAM variance is centred slightly south of the Equator at periods near 8.4 yr.

Separate AAM series were formed for the Northern Hemisphere subtropics (15.1°N – 31.4°N), the tropics (7.5°S – 7.5°N) and the Southern Hemisphere subtropics (31.4°S – 15.1°S). Variations of tropical AAM are significantly coherent with the subtropical AAM series for both hemispheres at periods centred on 4.2 yr (Fig. 1b) with variations in both the Northern and Southern Hemisphere subtropics tending to lag tropical AAM by about a quarter of a cycle for periods longer than ~ 2 yr (Fig. 1c).

To obtain a synoptic view of AAM variations, we formed a filtered time series from each latitude belt as the difference between centred 1-yr and 5-yr moving averages of the raw AAM values. This procedure efficiently removes all harmonics of the annual cycle, and has its maximum transfer function near a period of 4 yr, with half-power points at periods of ~ 2 and 8 yr. We combine the resulting interannual anomalies from the individual belts to form a latitude-time plot (Hovmöller diagram, Fig. 2a), in which warm colours (yellow to red) denote positive AAM anomalies (enhanced westerlies), and cool colours (blue to purple) indicate negative AAM anomalies (enhanced easterlies).

The diagram clearly shows the propagation of interannual AAM anomalies from the tropics to the subtropics, as inferred from the coherence plots (Fig. 1). Westerly anomalies in the tropics in 1981 and 1985 preceded strong El Niños by 2 yr, although the anomalies appearing in 1979 were not followed by an ENSO event (see later discussion). The westerly anomalies associated with the 1982–83 and 1986–87 events intensify as they propagate towards the subtropics, reaching their maximum strength during the mature phases of the corresponding El Niño episodes. The AAM anomalies then diminish in strength as they propagate towards the poles, with signals reaching high latitudes in both hemispheres up to five years after their origin in the tropics. The coherent propagation found over the northern

extratropics, in particular, makes it unlikely that these zonal wind anomalies are directly forced by fluctuations in the underlying sea surface temperature gradient associated with the ENSO cycle. The low-frequency variations may represent a combined mode of the coupled global ocean–atmosphere system, whose precise nature and origin remain to be determined.

In addition to strong variability at periods of 3–7 yr (hereafter referred to as the low-frequency (LF) component), the ENSO cycle has a distinct quasi-biennial (QB) component^{14–16}. We note here (Fig. 3a, b) that this bimodality is a feature of the records of both length of day and the Southern Oscillation Index over the past 15 years, with the length of day showing additional variability at ultra-low frequencies of non-atmospheric origin⁷. A spectrum of variations in the total AAM (Fig. 3c) during the

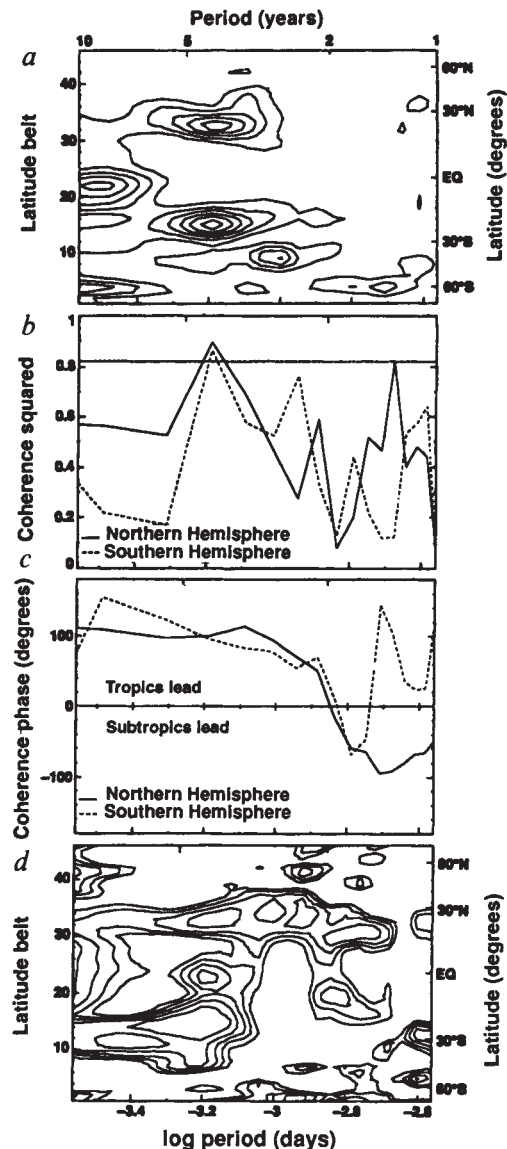


FIG. 1 a, Latitude-frequency dependence of observed AAM variance, as shown by a contour plot of power spectral density for the 46 equal-area belts numbered from south to north; the units are angular momentum squared $(1 \times 10^{22} \text{ kg m}^2 \text{ s}^{-1})^2$ per cycle per day; this plot uses a contour interval of 500, with contours starting at 750. A linear trend and the first and second annual harmonics were removed before computing the Fourier form. b, Coherence squared between tropical (from 7.5°N to 7.5°S) and subtropical (31.4° – 15.1°) AAM series, with 95% confidence level indicated. c, As in b, for the coherence phase. d, Latitude-frequency contour plot of the coherence squared between the Southern Oscillation index and the atmospheric angular-momentum data; this map uses a contour interval of 0.10, with contours starting at 0.50.

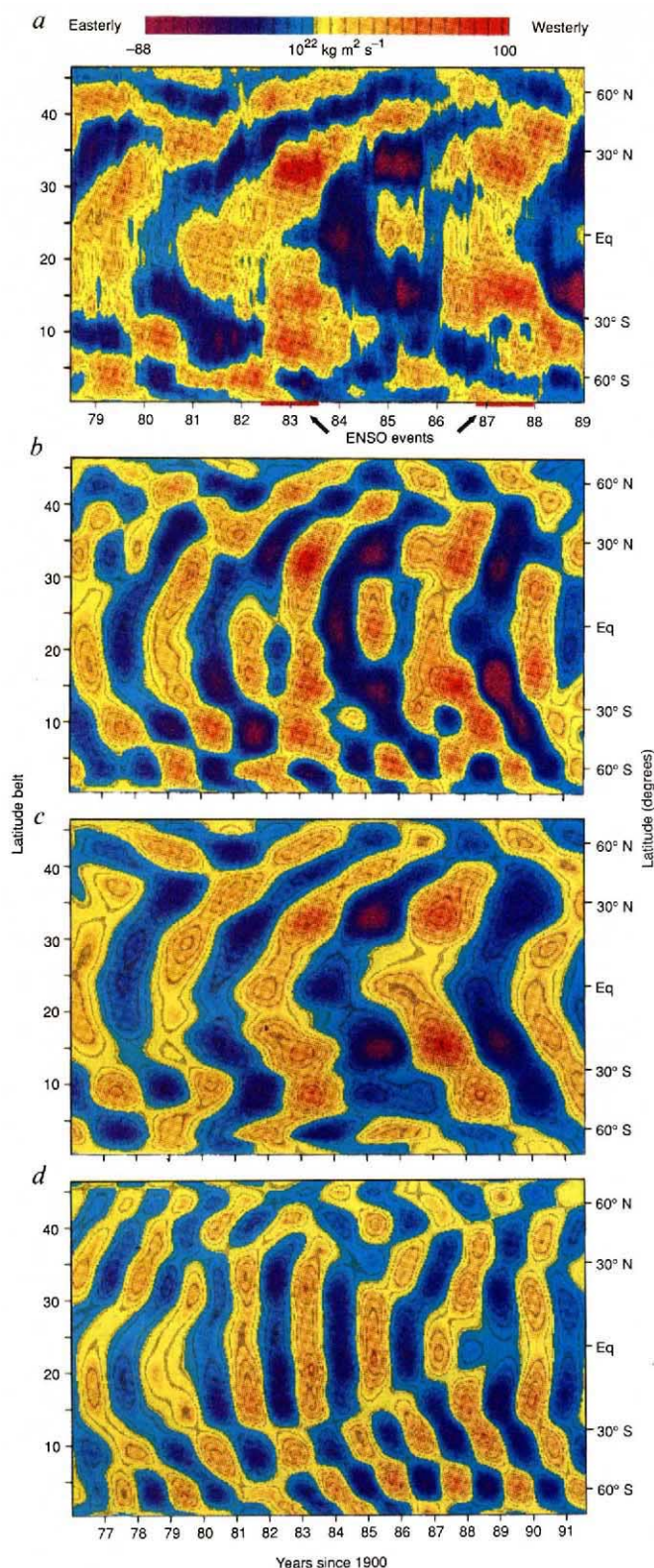


FIG. 2 Latitude-time (Hovmöller) plot of interannual atmospheric angular-momentum variations. We use the series of the National Meteorological Center⁵, integrated over 46 separate equal-area latitude bands based on atmospheric wind data up to 100 mbar¹⁸. *a*, Interannual variation, obtained from the difference between the one-year and five-year moving averages for each individual band. *b*, Interannual variation obtained by Fourier-filtering in the full interannual band (18–88 months); *c*, the low-frequency (LF) component obtained by Fourier-filtering in the 32–88-month band; *d*, the quasi-biennial (QB) component obtained by Fourier-filtering using an 18–35-month band.

same period shows a significant peak in the LF band, with no significant QB oscillation apparent. A QB oscillation signal in AAM integrated to the 50 mbar level has, however, been detected using adaptive filtering techniques¹⁷. Figure 1*d* shows the coherence of monthly AAM variations in individual latitude bands with the Southern Oscillation index; centres of strong coherence are evident in the NH and SH subtropics of both hemispheres and in the tropics in the LF band. Strong values are also found at the QB frequency ranging from 30° N to 25° S, related to coherent AAM variations in the subtropics and the tropics.

Prompted by the bimodality of these data types and the correlation of the Southern Oscillation with the banded AAM

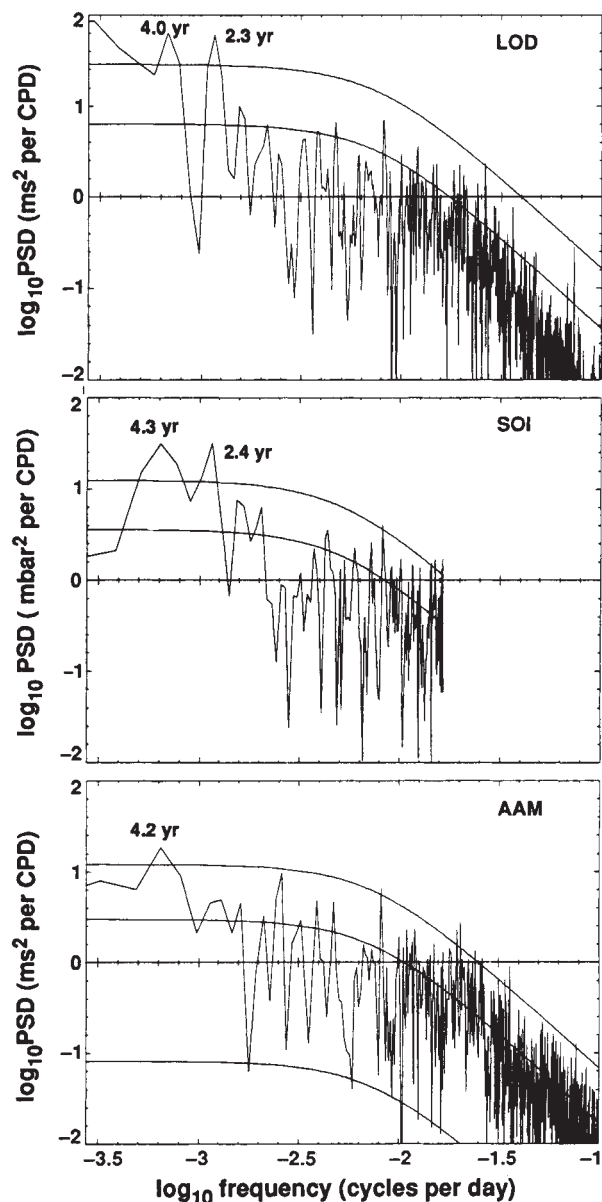
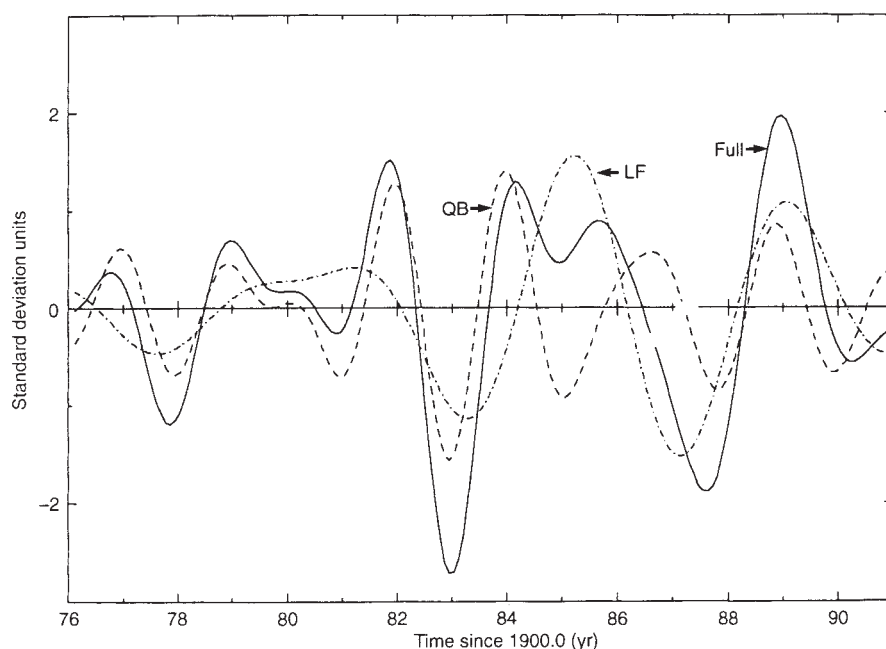


FIG. 3 Periodogram of length of day (LOD) as measured by space-geodetic techniques (top panel) and as inferred from the atmospheric angular-momentum (AAM) data summed over the global domain (bottom panel) where PSD indicates power spectral density and CPD denotes cycles per day, as well as that of the Southern Oscillation index (SOI, taken here as the difference in surface pressure between Tahiti and Darwin, middle panel), for the period 1976–91. A second-order polynomial, and the first and second annual harmonics, were removed from each series before computing the Fourier transform. Each series is compared to a first-order autoregressive (AR1) model (middle smooth curve) with 95% confidence levels shown (upper smooth curves). Significant peaks in the interannual range are indicated. The slight differences of corresponding periods between different data types result from different time intervals used in the three data sets.

FIG. 4 Fourier-filtered Southern Oscillation index: solid line indicates the interannual band (18–88 months); dashed line represents a quasibiennial filtered series (18–35 months); and the dashed-dotted line shows the low-frequency (LF) filtering (32–88 months).



at these periods, we applied Fourier filters to the Southern Oscillation Index series (Fig. 4) centred on three different time-scales: (1) a broad filter ('interannual band' from 18 to 88 months) encompassing both the QB and LF bands; (2) a 'LF' band for the period range 32–88 months; and (3) a 'QB' band focusing on the 18–35 month span. Similar filters were applied in ref. 16. ENSO events (in 1977, 1982–83 and 1987) occurred when the LF and QB components added constructively to produce a negative index with a magnitude of greater than one standard deviation. Conversely, a strong cold event took place in 1988 when the components constructively added to produce a significant positive index. 'Mild events' which do not meet the El Niño criteria (for example, 1980 and 1990) are associated with destructive interference between these two components. Similar behaviour has been found in sea surface temperature variations in the eastern equatorial Pacific¹⁴. Note that significant ENSO episodes appeared with alternate QB oscillation events: constructive interference gave ENSO events in 1977, 1982–83 and 1987; destructive interference occurred in 1980, 1985 and 1990, when events were absent. If this pattern holds true, one would expect an ENSO event during the next tropospheric QB oscillation.

The interannual Fourier filtering of the banded AAM shows the basic V-shaped pattern of poleward propagation (compare Figs 2a with Fig. 2b); the different behaviour of the LF (Fig. 2c) and QB (Fig. 2d) bands, however, suggests that the total ENSO signature results from interaction between two distinct components. The LF component has a striking poleward teleconnection (propagation) pattern, with events originating in the tropics and taking up to 5 yr to reach the polar regions. The teleconnection pattern of the QB band is more complex, with a standing wave visible in the tropical area and some indication of poleward propagation south of 25° S. North of ~20° N, propagation both towards the poles and towards the Equator is seen. The signal in the tropics is characterized by high coherence between AAM and the Southern Oscillation Index in this band (Fig. 1d), suggesting that the standing oscillations found in this region may be related to the QB component of the ENSO cycle.

Focusing on the 1982–83 and 1986–87 ENSO maxima, we note the complex interaction between the LF and QB components of the tropospheric AAM. The LF component shows a positive V-shaped pattern beginning at the Equator in early 1981, which reaches both the northern and southern subtropics two years later, at the height of the 1982–83 event. During this

time interval, the standing QB component has gone through one complete cycle, interfering constructively with the propagating LF component near the Equator in 1981 and in the subtropics of both hemispheres in 1983, similar to the QB component of the Southern Oscillation Index (negative anomalies in early 1981 and 1983). The modulation caused by the QB component is clearly visible as a tropical easterly anomaly (Fig. 2b) in 1982, although it is attenuated in Fig. 2a. A similar case applies for the 1986–87 ENSO; a positive LF teleconnection pattern emerges from the tropics in 1985, reaching the subtropics at the height of the event two years later, and interfering constructively with the QB component in 1985 and 1987; the strong negative modulation of the AAM by the QB component in late 1985 and early 1986 is clearly visible in the tropics (Figs 2a and b). Note that both the LF and QB bands showed strong negative (easterly) anomalies along the Equator in 1988 (Figs 2c and d); positive anomalies in the Southern Oscillation index (Fig. 4) in the LF and QB bands also added constructively at this time, resulting in an intense cold event (La Niña).

The later stages of the 1977–78 ENSO are evident in the subtropics for both the full interannual and LF component of the AAM (Fig. 2b and c), in agreement with Southern Oscillation Index (Fig. 4). Positive V-shaped AAM anomalies are also seen near the Equator in 1978 and 1989 propagating towards the pole and reaching the mid-latitudes in 1980 and 1991, respectively. Both the 1980 and 1990–91 'mild positive events' correspond to a net negative Southern Oscillation Index (Fig. 4). Therefore, it seems that all of the negative excursions of this index during the period of the study were associated with poleward-propagating positive (westerly) anomalies of interannual AAM, although not all cases are strong enough to be classified as full ENSO events. □

Received 6 December 1991; accepted 7 April 1992.

- Philander, S. G. H. *El Niño, La Niña and the Southern Oscillation* (Academic, New York, 1990).
- Barnes, R. T. H., Hide, R., White, A. A. & Wilson, C. A. *Proc. R. Soc. A* **387**, 31–73 (1983).
- Lambeck, K. *The Earth's Variable Rotation* (Cambridge Univ. Press, London, 1980).
- Rosen, R. D. & Salstein, D. A. *J. geophys. Res.* **88**, 5445–5451 (1983).
- Eubanks, T. M., Steppe, J. A., Dickey, J. O. & Callahan, P. S. *J. geophys. Res.* **90**, 5385–5404 (1985).
- Morgan, P. J., King, R. W. & Shapiro, I. I. *J. geophys. Res.* **90**, 12645–12652 (1985).
- Hide, R. & Dickey, J. *Science* **253**, 629–637 (1991).
- Chao, B. F. *Geophys. Res. Lett.* **11**, 541–544 (1984).
- Salstein, D. A. & Rosen, R. D. *J. Clim. appl. Meteorol.* **25**, 1870–1877 (1986).
- Eubanks, T. M., Steppe, J. A. & Dickey, J. O. in *Earth Rotation: Solved and Unsolved Problems* (ed. Cazenave, A.) 163–186 (Reidel, Hingham, Massachusetts 1986).
- Dickey, J. O., Eubanks, T. M. & Hide, R. *Variations in the Earth's Rotation* (eds McCarthy, D. D. & Carter, W. E.) 157–162 (American Geophysical Union, Washington DC, 1990).

12. Salstein, D. A. & Kann, D. M. in *International Earth Rotation Service, A. Rep.* 1990 (ed. Feissel, M.) 111-9-111-20 (Observatoire de Paris, 1991).
13. Yasunari, T. *J. met. Soc. Jpn* **67**, 483-493 (1989).
14. Rasmusson, E. V., Wang, X. & Ropelewski, C. F. *J. mar. Systems* **1**, 71-96 1990.
15. Keppenne, C. L. & Ghil, M. in *Proc. XVth A. Climate Diagnostics Workshop*, 30-35 (US Department of Commerce, NOAA, 1991).
16. Barnett, T. P. *J. Clim.* **4**, 269-288 (1991).
17. Penland, C., Ghil, M. & Weickmann, K. *J. geophys. Res.* **96**, 22659-22671 (1991).
18. Rosen, R. D., Salstein, D. A. & Wood, T. M. *J. geophys. Res.* **96**, 5145-5151 (1991).

ACKNOWLEDGEMENTS. We thank D. H. Boggs and C. L. Keppenne for the visualization of Fig. 2, D. L. T. Anderson, M. K. Davey, M. Ghil and R. D. Rosen for discussion and D. A. Salstein and R. D. Rosen for the banded AAM data. The work of J.O.D. and S.L.M. represents one phase of research done at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

Carbon isotope composition of atmospheric CO₂ during the last ice age from an Antarctic ice core

Markus Leuenberger*, Ulrich Siegenthaler* & Chester C. Langway†

* Physics Institute, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

† Department of Geology, State University of New York at Buffalo, Amherst, New York 14226, USA

BUBBLES of ancient air in polar ice cores have revealed that the atmospheric concentration of CO₂ during the Last Glacial Maximum was 180–200 p.p.m.v., substantially lower than the pre-industrial value of about 280 p.p.m.v. (refs 1, 2). It is generally thought that this reduction in atmospheric CO₂ during glacial time was driven by oceanic processes. The most likely explanations invoke either a decrease in dissolved CO₂ in surface waters because of a more efficient 'biological pump' transporting carbon to deep waters, or a higher alkalinity in the glacial ocean as a consequence of changes in carbonate dissolution or sedimentation³. Because isotope fractionation during photosynthesis depletes ¹³C in the organic matter produced, changes in the biological pump would alter the carbon isotope composition of atmospheric CO₂, whereas changes in alkalinity would in themselves have no such effect. Here we report measurements of the carbon isotope content of CO₂ in ice cores from Byrd Station, Antarctica, in an attempt to distinguish between these mechanisms. We find that during the ice age the reduced isotope ratio $\delta^{13}\text{C}$ was more negative than pre-industrial values by $0.3 \pm 0.2\%$. Although this result does not allow us to discriminate definitely between the two possible causes of lower glacial atmospheric CO₂, it does indicate that changes in the strength of the biological pump cannot alone have been responsible.

A flux of sinking organic particles continuously exports carbon from the ocean surface and leads to a depletion of total CO₂ (ΣCO_2) in surface water compared with the deep sea. It has been suggested⁴⁻⁸ that a change in this biological pump, which regulates the concentration of ΣCO_2 in surface water, was responsible for the lower glacial concentration of CO₂. A prime candidate region is the high-latitude ocean where the nutrient and ΣCO_2 concentrations, high at present, might have been reduced during glacial time because of less upwelling of CO₂-rich and nutrient-rich deep water, or stronger biological activity. Because $\delta^{13}\text{C}$ is lower in organic matter than in dissolved inorganic matter, the ΣCO_2 depletion in surface sea water accompanies a higher $\delta^{13}\text{C}$ than in deep-sea water. A more efficient biological carbon pump should therefore produce higher $\delta^{13}\text{C}$ in surface water and in atmospheric CO₂. Alternatively, the lower ice-age CO₂ might be explained by higher surface water alkalinity and therefore lower p_{CO_2} , due to a rearrangement of water masses^{9,10} or to changes in carbonate sedimentation; this would by itself not affect the $\delta^{13}\text{C}$ of atmospheric CO₂. Carbon isotope analyses therefore provide a means of discriminating between the different mechanisms.

We have measured $\delta^{13}\text{C}$ of CO₂ in air trapped in the deep ice core drilled at Byrd Station, Antarctica, in 1968 (ref. 11). The experimental method is an improved and simplified version of that used earlier^{12,13}. An ice sample (typically 500 g) is ground in vacuum with a milling cutter. From the air that escapes from the opened bubbles (extraction efficiency 75–85%), CO₂ is separated in the same vacuum system at liquid nitrogen temperature, and the remaining air is collected cryogenically at 20 K for $\delta^{15}\text{N}$ analysis. The CO₂ concentration is measured volumetrically. We determine $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ of the CO₂ samples, typically 9–13 μl at standard temperature and pressure, using a MAT 250 mass spectrometer. To correct the isotope ratios, we measure the N₂O/CO₂ ratio of the samples by mass spectrometry¹⁴. Extraction and analyses are tested regularly by processing an artificial air standard containing 272 p.p.m.v. of CO₂ together with artificial bubble-free ice. The whole procedure requires extreme care, as a very small amount of CO₂ gas must be processed in a relatively large system, and the risk of contamination is high.

The raw $\delta^{13}\text{C}$ results must be corrected for three effects. (1) N₂O, which is frozen out from the air samples along with the CO₂, has the same main molecular masses (44, 45, 46). The correction of the 45/44 mass ratio, corresponding to the N₂O/CO₂ ratio of the gas sample¹⁴, amounts to $0.24 \pm 0.02\%$ for the Byrd core samples; it happens to be the same for ice age and Holocene. N₂O results will be reported elsewhere (M.L. and U.S., manuscript in preparation). (2) We found that CO₂ extracted from standard air, processed together with a gas-free monocrystalline ice sample, has a $\delta^{13}\text{C}$ value that is systematically lower by $0.22 \pm 0.06\%$ than that of CO₂ extracted from a large standard air sample. We have not identified the cause of this extraction shift, but as it is reproducible, we have correspondingly corrected all $\delta^{13}\text{C}$ results. (3) We must correct for the gravitational separation of the gas components of different molecular weight in the firn overlying the ice sheet^{15,16}. The resulting enrichment of the heavy isotopes is, according to the barometric formula

$$\Delta\delta = \Delta Mzg/(RT) \quad (1)$$

where ΔM is the difference in molecular mass of the two molecules, g is the acceleration due to the gravity and z is the thickness of the firn layer in which the air components are in diffusive equilibrium. z is not initially known, so we measured $\delta^{15}\text{N}$ of N₂ for each extracted air sample; the gravitational $\delta^{13}\text{C}$ shift is, according to equation (1), equal to the measured $\delta^{15}\text{N}$ shift with respect to atmospheric N₂. The average $\delta^{15}\text{N}$ is 0.30‰ for the recent pre-industrial period, 0.36‰ for the early Holocene samples and 0.44‰ for the ice-age samples (standard deviation 0.05‰).

The $\delta^{13}\text{C}$ results and CO₂ concentrations are listed in Table 1 and plotted in Fig. 1, together with data from three different ice cores for the recent pre-industrial time. Our CO₂ values generally agree with earlier measurements^{1,17} obtained by a more precise method, although some of the oldest samples deviate by 10–15 p.p.m.v. The average ice-age $\delta^{13}\text{C}$ value is $-6.84 \pm 0.12\%$ (error is 2 standard deviations of the mean); a clear $\delta^{13}\text{C}$ trend for this period cannot be distinguished within the scatter of the data. For three results from old measurements¹², we cannot reliably reconstruct the extraction system correction. We do not consider these old results as reliable, and they are not taken into account in the mean. Two of the new results are clearly above the ice-age average, but we have no experimental reason to exclude them. Thus, we have to assume that the scatter of the ice-age $\delta^{13}\text{C}$ results is real. The early-Holocene average is $-6.65 \pm 0.13\%$. Four Holocene samples from the fractured, brittle-ice zone¹⁸ contained traces of drilling fluid which may have interfered with the isotope measurements (although we have no indication for this). If they are not included in the calculations, the early-Holocene average is -6.78% . The average value for the recent pre-industrial time is $-6.52 \pm 0.12\%$.

The best estimate from our experiments is that the atmospheric $\delta^{13}\text{C}$ in the period 20–40 kyr before present (BP) was $0.19 \pm 0.18\%$ more negative than in the early Holocene and $0.32 \pm 0.17\%$ lower than in the millenium preceding industrialization. The early-to-late Holocene $\delta^{13}\text{C}$ shift of 0.13% , although hardly significant within our uncertainty limits, might reflect the preferential withdrawal of isotopically light carbon during biomass buildup on land.

Isotopic changes in surface ocean water, as recorded in planktonic foraminifera, were summarized by Sarnthein and Winn¹⁹ who reported considerable regional scatter. The global average

$\delta^{13}\text{C}$ increase from glacial to Holocene may be 0 to 0.3%. If the temperature dependence of the air-sea $^{13}\text{C}/^{12}\text{C}$ fractionation is taken into account (see Table 1), the foraminiferal results correspond to a 0.2–0.5% lower atmospheric $\delta^{13}\text{C}$ during the ice-age, compatible with our ice core result.

Marino *et al.*²⁰ report, from measurements on C4 plants recovered from packrat middens, that the ice-age $\delta^{13}\text{C}$ value was 0.7–1.0% lower than during the Holocene. This is a larger change than we obtain. We see no reason why our ice-age isotope

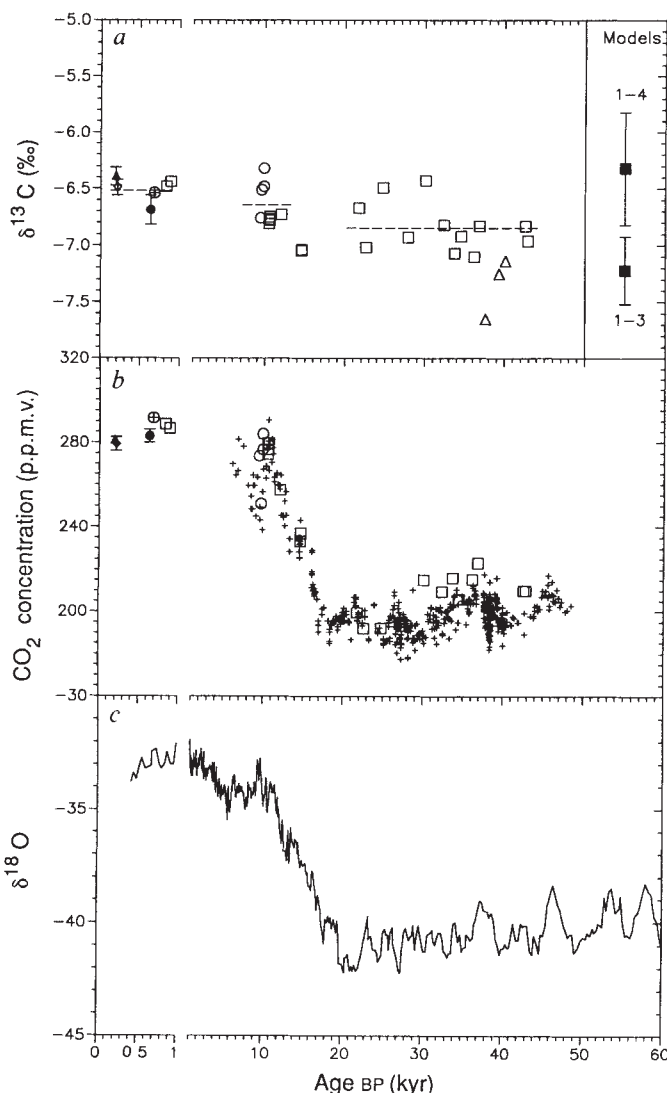


FIG. 1 *a*, $\delta^{13}\text{C}$ in atmospheric CO_2 . *b*, CO_2 concentrations, from ~40,000 BP to AD 1800, as measured on polar ice cores. $\square$, Byrd core (this work); $\circ$, Byrd core samples contaminated with drilling fluid; $\triangle$, Byrd core, old measurements¹²; $\diamond$, Siple station, Antarctica¹³ (average value); $\blacktriangle$, Dye 3, 1988 core (average value); $\bullet$, South Pole³¹ (average value); $\oplus$, new 1989 Byrd core³². Mean values are $-6.84 \pm 0.12\%$ (error is 2 standard deviations of the mean) for the ice age (excluding the three old, unreliable measurements ($\triangle$)), $-6.65 \pm 0.13\%$ for early Holocene and $-6.52 \pm 0.12\%$ for recent pre-industrial time. Dashed lines: average $\delta^{13}\text{C}$ values. CO_2 results for Byrd from refs 15, 16 are shown by crosses in *b*. Note that the CO_2 concentrations of Siple and Dye 3 can hardly be distinguished. On the right-hand side of *a*, '1–3' indicates the change expected from effects (1) to (3) listed in Table 2; '1–4' indicates the range for the sum of all effects of Table 2, including a change of the biological pump. The length of the error bar for '1–4' is somewhat arbitrary, as the uncertainty in the model results for a change in the biological pump is difficult to assess. *c*, Oxygen isotope record for the Byrd ice core.³³

TABLE 1 Carbon dioxide concentration and $\delta^{13}\text{C}$ in ice cores

Byrd core	Average depth (m)	Gas age (yr BP)	CO_2 concentration (p.p.m.v.)	$\delta^{13}\text{C}$ (‰)
	133.6	~680	291.4	-6.54*
	146.3	830	288.7	-6.48
	152.7	890	286.7	-6.44
	908.6	9,270	251.0	-6.51†
	928.9	9,550	274.0	-6.76†
	959.0	9,990	277.0	-6.48†
	959.1	9,990	284.0	-6.31†
	995.9	10,530	279.7	-6.81
	1,000.7	10,610	274.4	-6.77
	1,001.2	10,620	276.6	-6.75
	1,003.6	10,650	279.7	-6.78
	1,090.6	12,080	257.5	-6.73
	1,203.9	14,610	233.3	-7.05
	1,205.9	14,670	237.0	-7.04
	1,391.4	21,800	199.4	-6.67
	1,410.9	22,640	192.1	-7.02
	1,459.5	24,790	192.2	-6.49
	1,526.8	27,920	193.3	-6.93
	1,571.7	30,090	215.1	-6.43
	1,617.8	32,420	209.5	-6.82
	1,642.8	33,730	216.0	-7.07
	1,658.7	34,570	205.8	-6.92
	1,688.4	36,180	215.3	-7.10
	1,701.1	36,880	223.1	-6.83
	1,716.9	37,760	~217.0	-7.65‡
	1,754.6	39,930	~219.0	-7.25‡
	1,755.9	40,010	~209.0	-7.14‡
	1,798.9	42,590	210.0	-6.83
	1,803.8	42,890	209.9	-6.96

Average results for pre-industrial time

Ice core	Gas age (AD)	CO_2 (p.p.m.v.)	$\delta^{13}\text{C}$ (‰)	No. of results
Siple, Antarctica ¹³	1740–1820	279.3 ± 3.3	-6.50 ± 0.07	4
South Pole ³⁰	1200–1550	282.9 ± 3.1	-6.70 ± 0.13	9
Dye 3, Greenland§	1780–1820	281.1 ± 1.7	-6.41 ± 0.09	3
Byrd, Antarctica	1080–1310	288.9 ± 2.4	-6.49 ± 0.05	3

The estimated overall error (1σ) of the individual $\delta^{13}\text{C}$ analyses is $<0.10\%$. The dating of the Byrd core is subject to uncertainty¹⁷; the listed ages are calculated using a depth-age scale of C. U. Hammer (personal communication) for the ice and an air-to-ice age difference of 220 yr. Our Byrd $\delta^{13}\text{C}$ results as well as those for Dye 3 are corrected for N_2O , gravitational separation and an extraction shift. When published, the old Byrd results¹² and those for Siple¹³ and South Pole³¹ had not been corrected for these effects. The old results‡ are now corrected in total by $+0.24\%$, but this correction is uncertain; the results are shown for comparison only. We had to correct $\delta^{13}\text{C}$ for Siple in net by -0.10% (-0.32 for gravitational separation^{14,15}, equation (1), and 0.22% for the estimated extraction shift); for South Pole, the estimated correction is -0.37% ($= -0.59 + 0.22\%$), but its value is less certain than for the other cores.

* Youngest sample, from a new ice core drilled at Byrd station in 1989 (ref. 31).

† These four samples are contaminated with organic drill fluid, as identified from the mass spectrum; the CO_2 concentrations are therefore less precise, and the $\delta^{13}\text{C}$ values might be affected.

‡ Deep samples measured previously from Byrd¹² (the 'old' Byrd samples).

§ Dye 3 results are from a core drilled in 1988 (M.L., U.S. and J. Schwander, manuscript in preparation).

TABLE 2 Changes to atmospheric $\delta^{13}\text{C}$ between ice age and Holocene

Cause of change	Effect on glacial $\delta^{13}\text{C}$	Estimated contribution to $\delta^{13}\text{C}$ change
(1) Ice-age sea surface temperature $\sim 2^\circ\text{C}$ lower	Larger $^{13}\text{C}/^{12}\text{C}$ fractionation between sea-water ΣCO_2 and air- CO_2 (0.11‰ per $^\circ\text{C}$, ref. 35)	-0.2‰
(2) Change of land biomass or marine dissolved organic carbon	Whole-ocean $\delta^{13}\text{C}$ change	-0.32‰ (ref. 20)
(3) $\delta^{13}\text{C}$ of marine organic carbon $\sim 2\%$ higher during ice age ^{22,23}	Reduction of $\delta^{13}\text{C}$ difference between surface and deep sea by $\sim 10\%$	-0.2‰
(4) Change of the biological carbon pump (high-latitude nutrient hypothesis)	Increase of $\delta^{13}\text{C}$ difference between surface water and deep ocean	$\sim 0.9\%$ (model-dependent)

Factors that helped to produce the difference in $\delta^{13}\text{C}$ in atmospheric CO_2 from ice age to Holocene, and estimated resulting $\delta^{13}\text{C}$ changes (ice age minus Holocene). Effects (1) to (3) are independent of variations in the ocean circulation or the biological pump. The whole-ocean $\delta^{13}\text{C}$ change of 0.32‰, effect (2), implies a transfer of $\sim 500 \text{ Gt C}$ from the ocean to organic carbon with $\delta^{13}\text{C} \approx -25\%$, which is significantly less than the glacial to postglacial increase in terrestrial biomass of $1,300 \text{ Gt C}$ estimated by Adams *et al.*³⁴. The two estimates might be reconciled by a correspondingly higher ice-age pool of dissolved organic carbon in the oceans, but this possibility cannot be verified from observations. Effect (3) is based on the observation^{22,23} that $\delta^{13}\text{C}$ of organic carbon in glacial sediments is $\sim 2\%$ less negative than in Holocene sediments (for example, -18% instead of -20%), because there was $\sim 10\%$ less isotope fractionation between organic carbon and ΣCO_2 . The less-negative $\delta^{13}\text{C}$ of the organic particle rain must have led to a $\delta^{13}\text{C}$ difference between surface and deep water that was $\sim 0.2\%$, smaller than in pre-industrial time (2‰, ref. 10). The estimated uncertainty for the sum of effects (1) to (3) is 0.3‰. The results for models including high-latitude nutrients^{7,8}, effect (4), depend on the specific model assumptions (see text).

results should be systematically too high. On the other hand, it is conceivable that the isotope fractionation of the C4 plants varied under different climate conditions, although Marino *et al.* argue that this is unlikely. This discrepancy can only be resolved by further measurements.

What processes could explain our $\delta^{13}\text{C}$ results? In the absence of changes in ocean circulation or biology, three effects acting together would have depressed $\delta^{13}\text{C}$ of the ice-age atmosphere by $0.7 \pm 0.3\%$ (see Table 2). These include a whole-ocean isotope shift²¹, probably due to reduced biomass, the mentioned change in air-sea fractionation and a change in $\delta^{13}\text{C}$ of marine organic carbon^{22,23}. Our ice-core results show that the ice-age atmospheric $\delta^{13}\text{C}$ was probably higher than expected assuming a constant biological pump. This implies that the average oceanic surface-to-depth gradient in $\delta^{13}\text{C}$ (and therefore in nutrients) was larger than in the Holocene, an idea supported by the finding that the glacial minus Holocene $\delta^{13}\text{C}$ difference in planktonic foraminifera (0 to -0.3%) is less negative than in benthic foraminifera (-0.4 to -0.5% , ref. 24), as well as by higher glacial $\Delta\delta^{13}\text{C}$ (benthic-planktonic) values in individual sediment cores²⁴. Thus the carbon isotope evidence indicates a stronger biological pump during the ice age. For different model scenarios that invoke only changes in the biological pump (in the biological productivity or the vertical ocean circulation), the carbon-cycle box model of Wenk and Siegenthaler⁸ predicts an atmosphere-to-whole-ocean $\Delta\delta^{13}\text{C}$ change of 0.9‰ for a 90 p.p.m.v. lower concentration of atmospheric CO_2 (excluding effects (1) to (3) listed in Table 2); the ice-age scenario of Toggweiler and Sarmiento⁷, if linearly extrapolated to a 90 p.p.m.v. change in CO_2 , yields 1.0‰. (We consider a lowering by 90 p.p.m.v.; in fact oceanic CO_2 processes must have caused a larger change,

considering the reduced land biomass during the ice age; see Table 2.) These box model results would thus allow a slightly ($\sim 0.2\%$) higher glacial than postglacial atmospheric $\delta^{13}\text{C}$ for the high-latitude nutrient hypothesis, whereas our ice-core results yield a 0.3‰ lower glacial value (Fig. 1a). But there is not a one-to-one correspondence between changes of atmospheric CO_2 and $\delta^{13}\text{C}$, because nonbiological processes such as air-sea gas exchange or sea ice cover also influence the atmospheric $\delta^{13}\text{C}$. Recent studies based on multi-box models²⁶ and a three-dimensional ocean model²⁷ yield for some high-latitude nutrient scenarios a smaller atmospheric $\delta^{13}\text{C}$ change than the above simulations.

Sediment studies based on carbon isotopes in planktonic foraminifera^{28,29}, Cd/Ca ratios and accumulation rates of diatom shells³⁰ indicate that biological productivity was lower and nutrient levels in surface water were not reduced in the glacial Southern Ocean. This does not support the high-latitude nutrient hypothesis (but it does not necessarily contradict a stronger overall effect of the biological pump, for example if the Southern Ocean's surface area was reduced by sea ice). In summary, the combined experimental evidence from deep-sea sediments and the Byrd ice core, together with the models, suggests that the glacial-postglacial CO_2 variation was not caused by a change in the biological pump strength alone, although such a change may have contributed to it. A combination of changes in biological pump strength and in oceanic alkalinity^{3,9,10,20} could produce glacial-interglacial CO_2 variations that are compatible with our data and with deep-sea sediment records. A clear explanation of the glacial-interglacial CO_2 variations which is compatible with all existing observations is yet to be found. □

Received 24 January; accepted 20 May 1992.

- Neftel, A., Oeschger, H., Staffelbach, T. & Stauffer, E. *Nature* **331**, 609–611 (1988).
- Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **325**, 408–414 (1987).
- Broecker, W. S. & Peng, T.-H. in *NATO Volume on The Global Carbon Cycle* (ed. Heimann, M.) (in the press).
- Knox, F. & McElroy, M. B. *J. geophys. Res.* **89**, 4629–4637 (1985).
- Sarmiento, J. L. & Toggweiler, J. R. *Nature* **308**, 621–624 (1984).
- Siegenthaler, U. & Wenk, T. *Nature* **308**, 624–625 (1984).
- Toggweiler, J. R. & Sarmiento, J. L. in *Geophysical Monograph* 32 (eds Sundquist, E. & Broecker, W. S.) 153–184 (American Geophysical Union, 1985).
- Wenk, T. & Siegenthaler, U. in *Geophysical Monograph* 32 (eds Sundquist, E. & Broecker, W. S.) 185–194 (American Geophysical Union, 1985).
- Boyle, E. J. *geophys. Res.* **93**, 15701–15714 (1988).
- Broecker, W. S. & Peng, T.-H. *Global biogeochem. Cycles* **3**, 215–239 (1989).
- Ueda, H. T. & Garfield, D. *US Army Corps of Engineers, Tech. Rep.* 231 (Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, 1969).
- Friedli, H., Moor, E., Oeschger, H., Siegenthaler, U. & Stauffer, B. *Geophys. Res. Lett.* **11**, 1145–1148 (1984).
- Friedli, H., Löffler, H., Oeschger, H., Siegenthaler, U. & Stauffer, B. *Nature* **324**, 237–238 (1986).
- Friedli, H. & Siegenthaler, U. *Tellus* **B40**, 129–133 (1988).
- Schwander, J. in *The Environmental Records in Glaciers and Ice Sheets* (eds Oeschger, H. & Langway, C. C.) 53–68 (Wiley, Chichester, 1989).
- Craig, H., Horibe, Y. & Sowers, T. *Science* **2**, 1674–1678 (1988).
- Staffelbach, T., Stauffer, B., Sigg, A. & Oeschger, H. *Tellus* **B43**, 91–96 (1991).
- Shoji, H. & Langway, C. C. Jr in *Geophysical Monograph* 33 (eds Langway, C. C. *et al.*) 39–48 (American Geophysical Union, 1985).
- Sarnthein, M. & Winn, K. in *Climate–Ocean Interaction* (ed. Schlesinger, M.) 319–342 (Kluwer, Boston, Massachusetts, 1990).
- Marino, B. D., McElroy, M. B., Salawitch, R. J. & Spaulding, W. G. *Nature* **357**, 461–466 (1992).
- Duplessy, J. C. *et al. Paleoceanography* **3**, 343–360 (1988).
- Jasper, J. P. & Hayes, J. M. *Nature* **347**, 462–464 (1990).
- Rau, G. H., Froelich, P. N., Takahashi, T. & Des Marais, D. J. *Paleoceanography* **6**, 335–347 (1991).
- Sarnthein, M., Winn, K., Duplessy, J. C. & Fontugne, M. R. *Paleoceanography* **3**, 361–399 (1988).
- Shackleton, N. J., Hall, M. A., Line, J. & Cag, S. *Nature* **306**, 319–322 (1983).
- Keir, R. S. *Paleoceanography* **3**, 413–445 (1988); **5**, 253–276 (1990).
- Heinze, C., Maier-Reimer, E. & Winn, K. *Paleoceanography* **6**, 395–430 (1991).
- Labeyrie, L. & Duplessy, J. C. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **50**, 217–240 (1985).
- Charles, C. D. & Fairbanks, R. G. in *Geological History of the Polar Oceans: Arctic versus Antarctic* (eds Bleil, U. & Thiede, J.) 519–538 (Kluwer, Boston, Massachusetts, 1990).
- Mortlock, R. A. *et al. Nature* **351**, 220–223 (1991).
- Siegenthaler, U. *et al. Ann. Glaciol.* **10**, 151–156 (1988).
- Langway, C. C. & Osada, K. *Proc. Symp. Tropospheric Chemistry of the Antarctic Region* (Univ. of Colorado, Boulder, 1991).
- Johnson, S., Dansgaard, W., Clausen, H. B. & Langway, C. C. *Nature* **235**, 429–434 (1972).
- Adams, J. M., Faure, H., Faure-Denard, L., McGlade, J. M. & Woodward, F. I. *Nature* **348**, 711–714 (1990).
- Mook, W. G. *Neth. J. Sea Res.* **20**, 211–223 (1986).

ACKNOWLEDGEMENTS. We thank H. Oeschger, J. Schwander and B. Stauffer for advice and discussions, and R. S. Keir and D. Raynaud for comments. The laboratory research was supported by the Swiss NSF. The Byrd ice-core samples are part of a University of Bern and SUNY at Buffalo collaboration, supported by the U.S. NSF/DPP.

Proterozoic crustal extension and the generation of dome-and-keel structure in an Archaean granite-greenstone terrane

Stephen Marshak*, Fernando F. Alkmim†
& Hanna Jordt-Evangelista†

* Department of Geology, University of Illinois, 1301 West Green Street, Urbana, Illinois 61801, USA

† Departamento de Geologia, Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais 35.400, Brazil

ARCHAEOAN granite-greenstone terranes¹⁻³, in which narrow belts of 'greenstone' (ultramafic and mafic volcanics) and overlying sedimentary rocks occur in association with broad provinces of 'granite' (granitic igneous rocks, gneiss and migmatite), have long puzzled geologists because of the lack of any clear modern analogues. Not only is there uncertainty about how the greenstone formed in the first place⁴⁻⁶, but there is continuing debate about how and when the terranes developed their distinctive structural and metamorphic architecture. Many granite-greenstone terranes display a dome-and-keel geometry, in which belts of supracrustal (volcanic and sedimentary) rocks occur as structural troughs wedged between dome-shaped bodies of the granite-gneiss-migmatite complex. A metamorphic aureole typically occurs in the supracrustals adjacent to their contact with the domes, and the contact itself is a shear zone. Here we report the results of field studies of a granite-greenstone terrane in Brazil, showing that this architecture originated during the Proterozoic, more than 500 million years after extrusion of the greenstone. We suggest that the thermal regime and deformation kinematics necessary to

create this architecture could be generated during an episode of crustal extension, when hot basement rocks were transported upwards along a transcrustal normal fault system to the base of the supracrustals.

Although the geological features of granite-greenstone terranes vary widely, many share a characteristic architecture typified by the Quadrilátero Ferrífero (QF) region in Minas Gerais, Brazil. In this region⁷, which lies at the southern end of the São Francisco craton⁸ (Fig. 1a), a supracrustal assemblage composed of Archean greenstone (the Rio das Velhas supergroup, 2.7–2.9 × 10⁹ years (Gyr) old⁹) and Lower Proterozoic sedimentary rocks (the Minas supergroup, 2.1–2.4 Gyr old¹⁰) occurs in structural keels bordering domal bodies of Archean gneiss and migmatite (2.7–3.2 Gyr old; N. Machado and M. Carneiro, manuscript in preparation). Within the keels, gross stratigraphy defines tight synclinoria, or steeply dipping to overturned homoclines, but deformation in the keels is complex and primary stratification has been transposed locally into metamorphic foliation¹¹. Regionally, the supracrustals were metamorphosed to greenschist facies.

The eastern part of the QF region was deformed by west-verging thrusting during the 1.0–0.6 Gyr old Brasiliano event¹¹⁻¹³ (Fig. 1a), but in the western part of the region, older structures are preserved. West of the Brasiliano influence zone, rock on both sides of the contact between supracrustal assemblage and the granite-gneiss-migmatite complex has been strongly sheared. Next to these contacts, which dip away from the centres of the domes (Fig. 1b), units of the supracrustal assemblage are severely thinned or missing, and there is a metamorphic aureole 1 to 3.5 km wide. In a representative transect (Fig. 1c), pelitic units contain mineral assemblages of the cordierite-sillimanite zone immediately at the contact and decrease progressively in grade through the staurolite-andalusite-cordierite zone into the biotite zone away from the contact. Detailed petrography indicates that most metamorphism in the aureole occurred at a

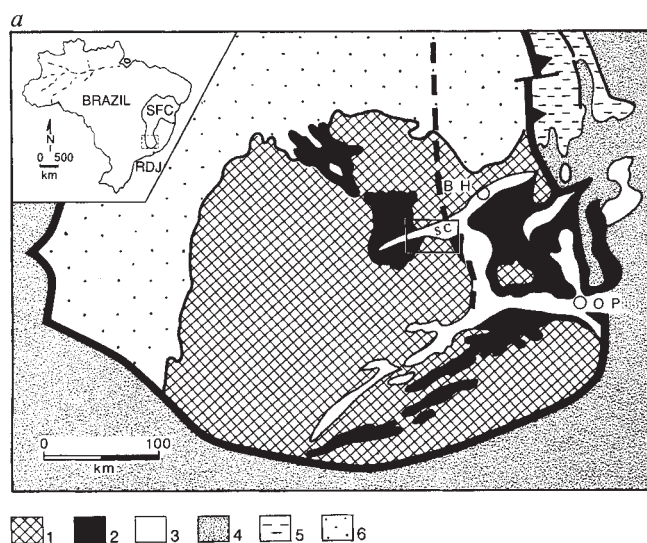
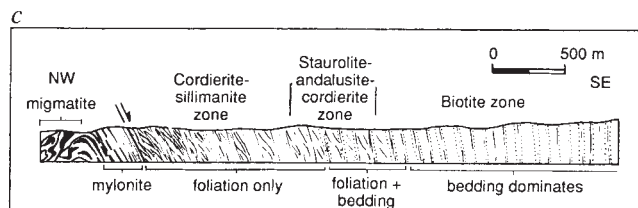
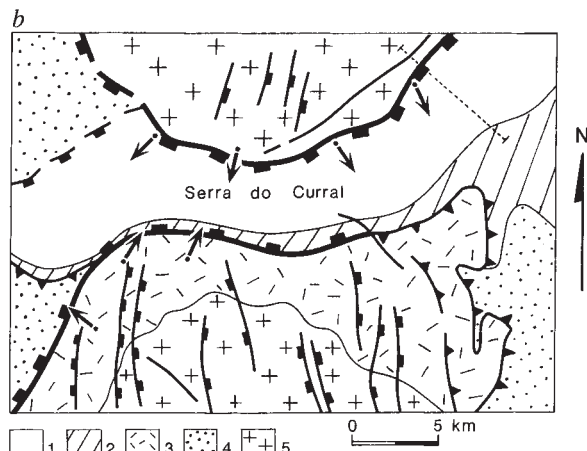


FIG. 1 Geological relationships in the Quadrilátero Ferrífero (QF) region. a, Geological sketch map of the southern end of the São Francisco craton in Brazil (location shown in the inset). The QF is the four-sided outcrop belt of supracrustal rocks. 1, Granite-gneiss-migmatite complex; 2, Rio das Velhas supergroup (greenstone sequence); 3, Minas supergroup (supracrustal sedimentary strata); 4, crust that was completely remobilized during the Brasiliano event; 5, Espinhaço supergroup; 6, unmetamorphosed Late Precambrian platform strata; SC, Serra do Curral; BH, Belo Horizonte; OP, Ouro Preto. The heavy dashed line shows the western limit of the Brasiliano influence zone. Lines with squared barbs are normal faults, and lines with triangular barbs are thrust faults. b, Map of a portion of the Serra do Curral (location is indicated by the box in a) based on our own mapping and on that of M. Carneiro and Dorr⁷. 1, Alberto Flores gneiss and migmatite; 2,



Souza Noschese granitoid; 3, Rio Das Velhas supergroup; 4, older units in the Minas supergroup; 5, younger units in the Minas supergroup. The arrows indicate mineral lineations in contact zone mylonites. c, Field relationships along traverse XX', showing metamorphic zonation and distribution of structural fabrics. Location of the traverse is shown by the thin dashed line in the northeast corner of b.

pressure of ~ 300 MPa and that the aureole reached peak temperatures of $\sim 600^\circ\text{C}$ (ref. 14). Locally, lenses of granitoid occur just below the contact (Fig. 1b). This granitoid has igneous textures overprinted by mylonitic shear fabric, but it does not contain compositional banding characteristic of gneiss and migmatite in the interior of the domes. Where the granitoid is absent, Archean gneiss is juxtaposed against supracrustal rocks. For example, on the north flank of the Serra do Curral, west of Belo Horizonte, sheared gneiss is in direct contact with the top unit of the Minas supergroup. Analysis of fabrics and kinematic indicators in rocks of the contact zone indicates that porphyroblast growth in the metamorphic aureole was syntectonic to post-tectonic with respect to development of the shear fabrics, and that shear was normal in sense (the upper 'hanging wall' having moved downwards) and radial with respect to the centres of the domes.

Cross-cutting relations allow us to place the radial hanging-wall-downwards shear and the metamorphism into a relative chronology. The structural juxtaposition of the granite-gneiss-migmatite complex against Lower Proterozoic (Minas) meta-sedimentary rocks and the formation of the associated metamorphic aureole indicates that the development of the dome-and-keel geometry is Lower Proterozoic or younger. In other words, this architecture is at least 500 million years younger than the extrusion of the greenstone. This proposal is further supported by the observation that there is no relation between the geometry of the keels and the distribution of sedimentary facies in the supracrustal sequence. Extensional shear fabrics of the contact zone overprint the foliation of Archean gneiss, as well as igneous textures of the granitoid. Fabrics and metamorphism of the contact zone also overprint early regional lower greenschist-facies metamorphism and an early foliation in the supracrustal sequence, indicating that supracrustal rocks had already been deformed and metamorphosed before creation of the domes and keels. We suggested previously¹¹ that this early metamorphism and deformation of the supracrustals was a consequence of the ~ 2.0 -Gyr-old Transamazonian orogeny. If this premise is correct, then the dome-and-keel geometry and the metamorphic aureole are late or post-Transamazonian, a proposal that is compatible with K-Ar dates of 1.7 Gyr on the uplift of the granite-gneiss-migmatite complex. The rise of the basement domes occurred before the Brasiliano event, because dome-and-keel geometry controls the geometry of Brasiliano thrusts and folds¹³.

A model to explain the architecture of the granite-greenstone terrane in the QF, and of similar ones globally, must explain how a gneiss-migmatite complex that was hot enough to create a metamorphic aureole was juxtaposed against relatively cool supracrustal rocks long after the supracrustal assemblage was created. We suggest that such a thermal regime could be established during regional crustal extension involving displacement along a transcrustal normal fault system¹⁵. Movement on such a system could bring relatively hot basement rock from deep crustal levels up to the base of the hanging-wall supracrustal assemblage. Because rock in the footwall is moving past the hanging wall, the heat source is continuously replenished, and thus even if the footwall temperature were only ~ 650 – 700°C , the observed metamorphic temperatures of hanging-wall supracrustal rocks could be achieved (K. Stüwe, unpublished computer model). The high initial temperature of the deep crust of the footwall could have been caused by thinning of the lithosphere and advection of heat into the base of the crust during rift-associated basaltic underplating^{16–18}. The introduction of water, released from the supracrustals by dehydration reactions, aided by depressurization, possibly stimulated partial melting of footwall rocks and the formation of granitoid lenses. Occurrence of discrete mylonitic shear zones in the contact zone suggests that the final stage of the displacement occurred after the rocks had cooled below peak metamorphic temperatures.

An extensional model is not the only way to introduce heat to the base of the predeformed supracrustal assemblage. Other possibilities include shear heating during thrusting¹⁹ and advection of heat into the basement by extensive intrusion^{20–23}. This post-tectonic igneous activity could have been caused by delamination of thickened lithosphere mantle^{24,25} after the Transamazonian orogeny. At present we favour the extensional model because there is abundant independent evidence for extensional tectonism after the Transamazonian and before the Brasiliano in the QF region¹¹, and we do not see field evidence for extensive post-tectonic igneous activity associated with the formation of the domes (the domes are not cored by granitoid but are composed of old gneiss; nonfoliated granitoid occurs only along the margins of the domes, and there are no dykes or pods of granitoid cutting the supracrustals).

The final rise of the domes may have been due to unloading²⁶, or it may have been a diapiric process triggered by a density inversion established when relatively hot and ductile intermediate-composition gneiss was juxtaposed beneath cooler and

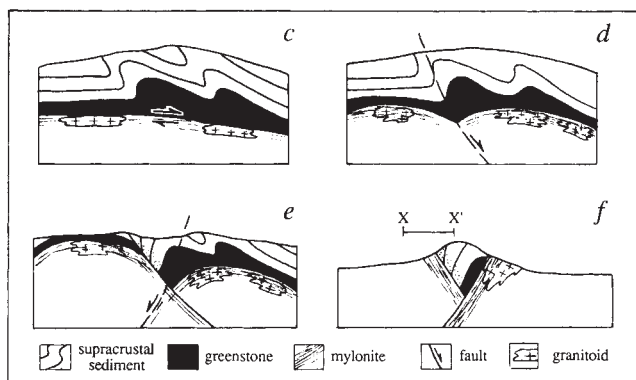
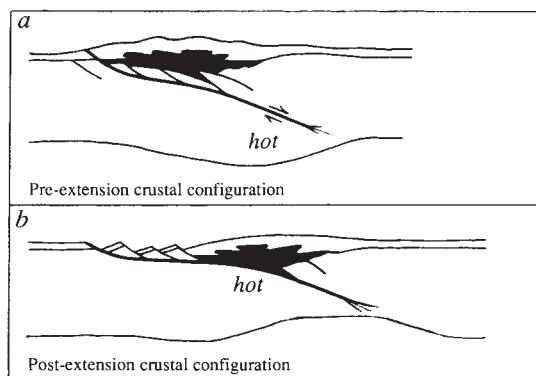


FIG. 2 Cross-sectional sketches illustrating an extensional model for the evolution of granite-greenstone terranes. *a*, Schematic crustal section subsequent to the closure of the greenstone basin during the Transamazonian orogeny but before transcrustal extension; *b*, schematic crustal section after displacement on a transcrustal normal fault system, emphasizing the juxtaposition of hot basement against cooler supracrustal rocks; *c*, detail showing the configuration of the basement and supracrustal contact region after hot deep crust has been juxtaposed against the base of the greenstone

sequence, metamorphism has begun, and granitoid has formed. *d*, Initiation of basement-cored domes and the association of dome boundaries with normal faults; *e*, continued rise of the basement and initiation of a second normal fault; *f*, present configuration of domes and a keel, such as is now exposed at the Serra do Curral. Dotted region shows metamorphic aureole in supracrustal sediment (metamorphic aureole is also present in the 'greenstone' at the contact). Traverse XX' is shown in Fig. 1c.

denser mafic or ultramafic igneous rocks and iron-rich meta-sedimentary rocks^{21,27}. Diapirism is implied by the radial pattern of shear around the domes. The density inversion leading to diapirism could have formed regardless of the mechanism that introduced heat to the QF region. As the domes rose, predeformed supracrustal rocks dropped into shear-zone-bounded slots between the domes. The boundaries of these slots presumably extended upward as normal faults (Fig. 2c). Metamorphism may have continued after the tectonic activity because of heat focusing by the domes themselves²⁸.

The extensional event proposed for the QF region could be related to the development of the Espinhaço basin which is preserved to the northeast of the QF region (Fig. 1a); the occurrence of extensional shear zones trending north-south in the basement of the QF region (Fig. 1) and of north-south dykes throughout the southern São Francisco craton suggests that extension occurred in an east-west direction, a strain axis that is also compatible with opening of the Espinhaço basin¹. The extensional model implies that the high-grade gneiss now next to the base of the supracrustal assemblage in the QF was not the basement on which the supracrustal assemblage was originally deposited. Similar allochthoneity characterizes the Cenozoic core-complex terranes of the western United States¹⁸. Differences in structural style between the QF region and Cenozoic core-complex terranes may reflect higher geothermal gradients, the density contrast between the hanging wall and footwall, and/or differences in the level of exposure. Documentation of extensional events similar to that of the QF region elsewhere in the world may provide a constraint on the chronology of proposed supercontinent cycles in the Proterozoic. Although an extensional model might not explain the architecture of all granite-greenstone terranes (see alternative explanations in refs 1, 2, 29–32), its applicability should be tested by examining kinematics of movement on shear zones³³, studying the chemistry and geochronology of related granitoids and metamorphic rocks, and developing thermal models. □

Received 27 December 1991; accepted 14 April 1992.

- Anhaeusser, C. R., Mason, R., Viljoen, M. J. & Viljoen, R. P. *Geol. Soc. Am. Bull.* **80**, 2175–2200 (1969).
- Goodwin, A. M. *Precambrian Geology: The Dynamic Evolution of the Continental Crust* (Academic, London, 1991).
- Platt, J. P. *Tectonophysics* **65**, 127–150 (1980).
- Boutter, C. A., Bickle, M. J., Gibson, B., & Wright, R. K. *Precamb. Res.* **36**, 241–258 (1987).
- Swager, C. & Griffin, T. J. *Precamb. Res.* **48**, 63–73 (1990).
- Kröner, A. in *Precambrian Plate Tectonics* (ed. Kröner, A.) 57–90 (Elsevier, Amsterdam, 1981).
- Dorr, J. V. N. II *U.S. Geol. Surv. Prof. Pap.* **641-A**, 110 (1969).
- Almeida, F. F. M. *Rev. Bras. Geoci.* **7**, 349–364 (1977).
- Machado, N., Noce, C. M., Oliveira, O. A. B. & Ladeira, E. A. *Bol. Soc. Brasil. Geoci. Núcleo Minas Gerais* **8**, 1–5 (1989).
- Babinsky, M., Chemale, F. Jr & Van Schmus, W. R. *Bol. Resum. 3rd Congr. Bras. Geoquímica*, 628–629 (Socied. Brasil. Geol., São Paulo, 1991).
- Marshak, S. & Alkmim, F. F. *Tectonics* **8**, 555–571 (1989).
- Almeida, F. F. M. & Hasui, Y. in *O Precambriano do Brasil* (eds Almeida, F. F. M. & Hasui, Y.) 1–5 (Blücher, São Paulo, 1984).
- Chemale, F. Jr, Rosiere, C. A. & Endo, I. *Rev. Pesquisas Univ. Fed. Rio Grande do Sul* (Univ. Fed. Rio Grande do Sul, Porto Alegre, in the press).
- Jordt-Evangelista, H., Alkmim, F. F. & Marshak, S. *6th Simp. Geol. Minas Gerais* (Revista Escola de Minas (REM), Ouro Preto, in the press).
- Marshak, S., Alkmim, F. F. & Evangelista, H. *Geol. Soc. Am. Abstr. Programs* **23**, 308 (1991).
- Sandiford, M. & Powell, R. *Earth planet. Sci. Lett.* **79**, 151–158 (1986).
- White, R. S. & McKenzie, D. P. *J. geophys. Res.* **94**, 7685–7730 (1989).
- Lister, G. S., Etheridge, M. A. & Symonds, P. A. *Tectonics* **10**, 1039–1064 (1991).
- Molnar, P. & England, P. *J. geophys. Res.* **95**, 4833–4856 (1990).
- Cowie, K. C. *Plate Tectonics & Crustal Evolution*, 3rd edn (Pergamon, Oxford, 1989).
- Schwerdtner, W. M., Morgan, J. & Stott, G. M. *Geol. Assoc. Can. Spec. Paper* **28**, 117–123 (1985).
- Fyson, W. K., Herd, R. K. & Ermanovics, I. F. *Can. J. Earth. Sci.* **15**, 1817–1825 (1978).
- Collins, W. J. & Vernon, R. H. *Geology* **19**, 835–838 (1991).
- Houseman, G. A., McKenzie, D. P. & Molnar, P. *J. geophys. Res.* **86**, 6115–6132 (1981).
- Turner, S. & Foden, J. *Geol. Soc. Austr. Abstr.* **25**, 262–263 (1990).
- Spencer, J. E. *Geology* **12**, 95–98 (1984).
- Park, R. G. *Geol. Rund.* **71**, 22–37 (1982).
- Allen, T. & Chamberlain, C. P. *Earth planet. Sci. Lett.* **93**, 392–404 (1989).
- Gorman, B. E., Pearce, T. H. & Birkett, T. C. *Precamb. Res.* **6**, 23–41 (1978).
- Spry, J. G. *J. struct. Geol.* **7**, 187–203 (1985).
- Hudleston, P. J., Schultz-Ela, D. & Southwick, D. L. *Can. J. Earth Sci.* **25**, 1060–1068 (1988).
- Feng, R. & Kerrich, R. *Geology* **18**, 870–873 (1990).
- Nalini, E. A. Jr & Hippert, J. F. *6th Simp. Geol. Minas Gerais* (Revista Escola de Minas (REM), Ouro Preto, in the press).

ACKNOWLEDGEMENTS. We thank the reviewers, P. Hoffman and A. Hynes, for comments, and K. Stüwe, M. Sandiford, J. Foden, P. James, M. Carneiro and H. Brueckner for suggestions.

Parasitism in a 400-million-year-old green alga

T. N. Taylor*, W. Remy† & H. Hass†

* Department of Plant Biology and Byrd Polar Research Center, Ohio State University, Columbus, Ohio 43210, USA

† Westfälische Wilhelms-Universität, Forschungsstelle für Paläobotanik, Hindenburgplatz 57–59, D-4400 Münster, Germany

SEVERAL stages of a parasitic fungus were found in the cells of the alga *Palaeonitella* in thin sections of macroplants in the Lower Devonian Rhynie chert. Some of the algal cells are greatly enlarged indicating a host response. The fungi include encysted zoospores on the algal cell wall, penetration of the host cell wall, and transfer of the fungal protoplast and the mature zoosporangium with discharge tube. Here we document evidence of plasmodiophoromycetes in the fossil record and demonstrate that this type of parasitic interaction with a modern host response was well established in freshwater palaeoecosystems at least 400 million years ago.

One of the most remarkable fossil assemblages is from the Lower Devonian (Siegenian) Rhynie Chert in Aberdeenshire, Scotland¹. This site, now known to have been a freshwater palaeoecosystem, has provided detailed reports of permineralized macroplants of Devonian age, and has contributed information ranging from new theories about the structure of conducting elements² in the early land plants, to the first report of their gametophyte generation^{3,4}. It has become apparent that a diverse mycoflora existed with the Rhynie chert plants (T.N.T., W.R. and H.H., manuscript in preparation). Some of these fungal-plant interactions were saprophytic, whereas others were clearly parasitic.

Within the Rhynie chert are several stages in the life history of a plasmodiophoromycetous fungus that was parasitic on the charophyte *Palaeonitella* (Fig. 1a)^{1,5}. This alga consists of branched, septate filaments that are attached at distinct nodes, much like the extant charophyte *Nitella*. Normal cells of *Palaeonitella* range from 30 to 70 µm in diameter (Fig. 1a) and may contain irregularly-shaped, dense granules (Fig. 1b). On the surface of some cells are spherical, thick-walled, encysted zoospores. They range from 3 to 5 µm in diameter, and often appear in slight depressions along the cell wall. In some sections there is a narrow line extending through the cell wall that represents the infection tube. In other cells the empty zoospore cyst can be seen on the outside of the host cell (Fig. 1c), whereas inside, immediately beneath the zoospore cyst, is a dense spherical structure that represents the fungal protoplast (Fig. 1c). It is from this protoplast that the zoosporangium would subsequently have developed. Other *Palaeonitella* cells contain flask-shaped structures that represent mature plasmodiophoromycete zoosporangia (Fig. 1d). They are up to 30 µm long and consist of a swollen base and narrow neck. In some forms the zoosporangia are irregular in outline and flaccid suggesting that zoospores were discharged before fossilization. Each zoosporangium is characterized by a single, narrow (6–10 µm long) discharge papilla that always penetrates the algal cell wall (Fig. 1b). In some specimens a slight thickening is present near the orifice of the neck (Fig. 1b).

In the absence of most stages in the life history of the fossil fungus, as well as certain fine ultrastructural details, it is impossible to determine the precise systematic position of these endobiotic parasites. Despite this, however, the morphology and size of the fossil zoosporangia are similar to several plasmodiophoromycetous taxa reported by Karling (for example *Woronina* and *Sorodiscus*)⁶.

Documenting the level of interaction between a fungus and its host is especially challenging in fossil specimens. In the absence of an obvious host response, parasitism may be indistinguishable from saprophytism. In the case of these Devonian

fungi, however, the parasitic interaction between the alga and fungus is substantiated by the distinct hypertrophy of some of the internodal cells of *Palaeonitella*. Whereas the normal cell diameters of *Palaeonitella* range from 30 to 70 μm , the infected hypertrophied cells of the alga are in the 200–300 μm range. The presence of these enormous cells intermixed with normal ones is exactly the same host response that has been reported in two living species of *Chara* (*C. contraria* and *C. delicatula*)^{6,7}. The parasitic fungus responsible for this hypertrophied condition is the plasmodiophoromycete *Sorodiscus*. In this modern example of endoparasitism, several stages in the life history (for example sporangial plasmodium, sporangiosori, sporangia and zoospores) are unknown. In addition to extensive swelling of

the infected cells, the modern parasites are believed to be responsible for the production of storage starch in the plastids⁷. Perhaps some of the dark granular material present in the cells of *Palaeonitella* (Fig. 1b) signal this response in the fossil as well.

Although several groups of fungi have been documented from the fossil record, their phylogeny continues to remain problematic. Nevertheless, the presence of several different types of aquatic fungi in the Rhynie chert community underscores the fact that as a group these fungi had already established numerous levels of interaction by Lower Devonian time. What is perhaps most surprising is that the host symptoms in at least one of these interactions has remained unchanged for at least 400 million years. □

Received 17 February; accepted 15 April 1992.

1. Kidston, W. H. & Lang, W. H. *Trans. R. Soc. Edinb.* **52**, 855–902 (1921).
2. Edwards, D. S. *Bot. J. Linnean Soc.* **93**, 173–204 (1986).
3. Remy, W. & Remy, R. *Argument. Palaeobot.* **6**, 37–72 (1980).
4. Remy, W. & Hass, H. *Argument. Palaeobot.* **8**, 69–117 (1991).
5. Edwards, D. S. & Lyon, A. G. *Bot. J. Linnean Soc.* **86**, 37–55 (1983).
6. Karling, J. S. *The Plasmodiophorales* 1–256 (Hafner, New York, 1968).
7. Karling, J. S. *Am. J. Bot.* **15**, 485–496 (1928).

ACKNOWLEDGEMENTS. This work was supported by the National Science Foundation, Deutsche Forschungsgemeinschaft (Re 200/16-1) and the Deutscher Akademischer Austausch Dienst.

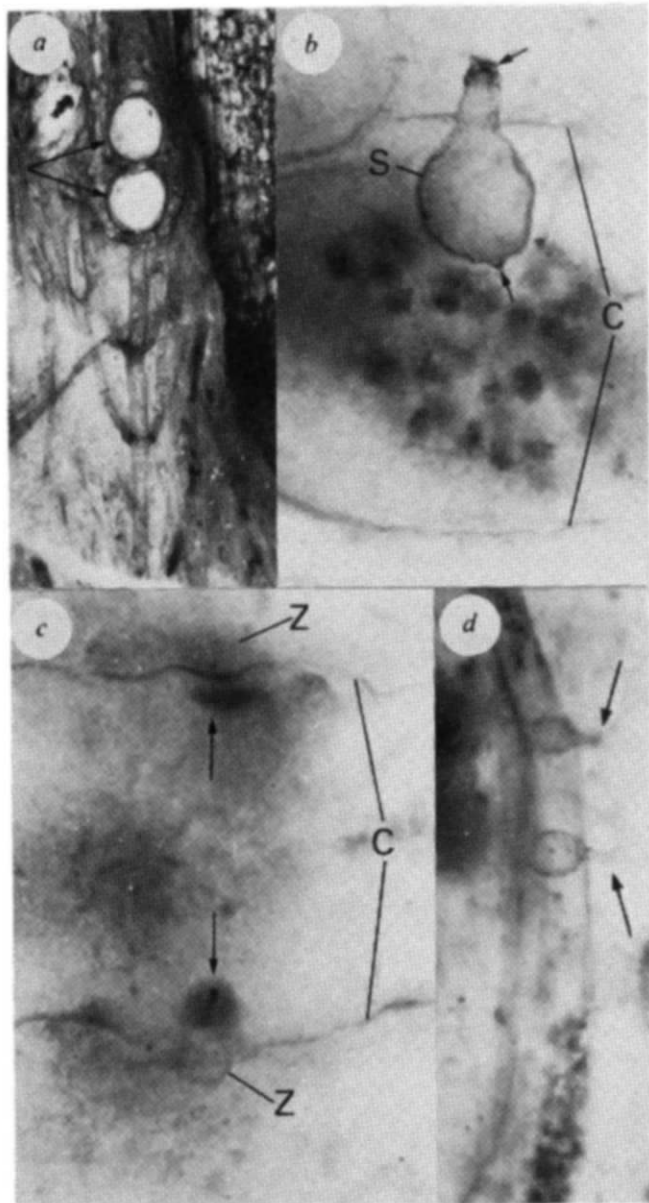


FIG. 1 Lower Devonian fungi. *a*, Longitudinal section showing several nodes of *Palaeonitella* sp. Arrows indicate hypertrophied internodal cells. Magnification $\times 30$. *b*, Zoosporangium (S) with discharge papilla extending through *Palaeonitella* cell wall (C). Upper arrow indicates thickening near orifice; inner arrow marks thickening at base of zoosporangium. Magnification $\times 1,000$. *c*, Section through cell (C) with two encysted zoospores (Z) on outside and protoplasts (arrows) inside. Magnification $\times 1,500$. *d*, *Palaeonitella* cell with two zoosporangia. Note discharge papillae extending through cell wall (arrows). Magnification $\times 300$.

Extra-pair paternity results from female preference for high-quality males in the blue tit

Bart Kempenaers, Geert R. Verheyen, Marleen Van den Broeck*, Terry Burke†, Christine Van Broeckhoven* & André A. Dhondt

Department of Biology and * Department of Biochemistry, University of Antwerp, Universitaire Instelling Antwerpen B-2610, Wilrijk, Belgium

† Department of Zoology, University of Leicester, LE1 7RH, UK

EXTRA-PAIR copulations (EPCs) seem to be one of the most widespread alternative reproductive behaviours by which male birds can increase their fitness^{1,2}. In many species females actively solicit or freely engage in EPCs^{3–5}, which suggests that they benefit from them. Of the eight hypothetical benefits proposed^{2,6}, the most likely are genetic². Often females engage in EPCs with more dominant males^{3,7} or with males with more elaborate ornaments^{8,9}. In species in which paternity was assigned, extra-pair young were divided asymmetrically between males^{10–12}. Here, combining detailed behavioural work with DNA-fingerprinting of an entire population, we present evidence that such an asymmetry is indeed caused by female behaviour, and that 'attractive' males do not suffer lost paternity, survive better and recruit more young. Our results support the genetic quality hypothesis.

We have studied a colour-ringed blue tit *Parus caeruleus* population in plot C (an isolated 17 ha wooded estate north of Antwerp, Belgium; ref. 13) since 1979. Most males are monogamous, but some have two or (rarely) three females¹⁴. In 1990, 39 females bred with 32 males, in 1991, 43 bred with 38 males. Mate guarding behaviour was studied in detail for 14 nests in 1990 (eight monogamous pairs and six females of polygynous males) and for 31 nests in 1991 (20 monogamous pairs and 11 females of polygynous males) (Table 1). All adults were trapped and length of wing and tarsus was measured. We collected blood samples (10–50 μl) for DNA fingerprinting, by puncture of the brachial vein, from all adults and from all 14-day-old nestlings in 1990. In addition we collected two nests with nestlings that had died earlier.

The DNA fingerprint analysis (Fig. 1: multi-locus) shows that in 11 of 36 nests (31%) at least one excluded offspring is present. Of 314 nestlings, 33 are extra-pair young (EPY; 11%). In the 11 nests with EPYs 33 of 90 young (37%) were fathered by an extra-pair male. If EPY are divided randomly over the nests, one expects them to be Poisson distributed, under the assumption that the occurrence of one extra-pair fertilization is independent of previous occurrences. This assumption is probably realistic, because (1) one egg is fertilized each day², (2) copulation frequency is high before and during egg-laying (0.5 to 2 copulations per hour throughout the day; our unpublished observation), and (3) last mate sperm precedence is likely^{2,15}. A loglinear model¹⁶, controlling for number of young per nest, shows that EPY are not randomly distributed over nests ($X^2 = 79.32$, d.f. = 34, $P < 0.05$). This can occur because males differ in mate guarding ability, or because some males but not others try to obtain EPCs, or because of differences in female behaviour.

Although monogamous males guard their mates more closely than polygynous males, they do not have fewer EPY in their nests (Table 1). Thus, mate guarding seems to be relatively ineffective in protecting paternity. Over two years we observed 71 copulations, of which seven (10%) were EPCs, all involving a neighbouring male. Five EPCs occurred on the territory of the neighbour. In the other two cases the female engaged in a copulation with an intruding neighbour. During the fertile period, we observed 277 male and 206 female intrusions, of which 95% and 85%, respectively, were intrusions by known neighbours. The intruding female was either aggressively chased away by the resident female (53% of 175 observations) or approached by the resident male. After the male approach the female either left the territory (38%), showed the typical wing-shiver display before copulation (6%) or copulated with the male (3%). These data, combined with the observation that females never foraged during their intrusions, nor received any courtship feeding, suggest that intruding females are seeking EPCs.

Female intrusion rates differed markedly between territories

TABLE 1 Mate guarding and extra-pair paternity for monogamous and polygynous broods

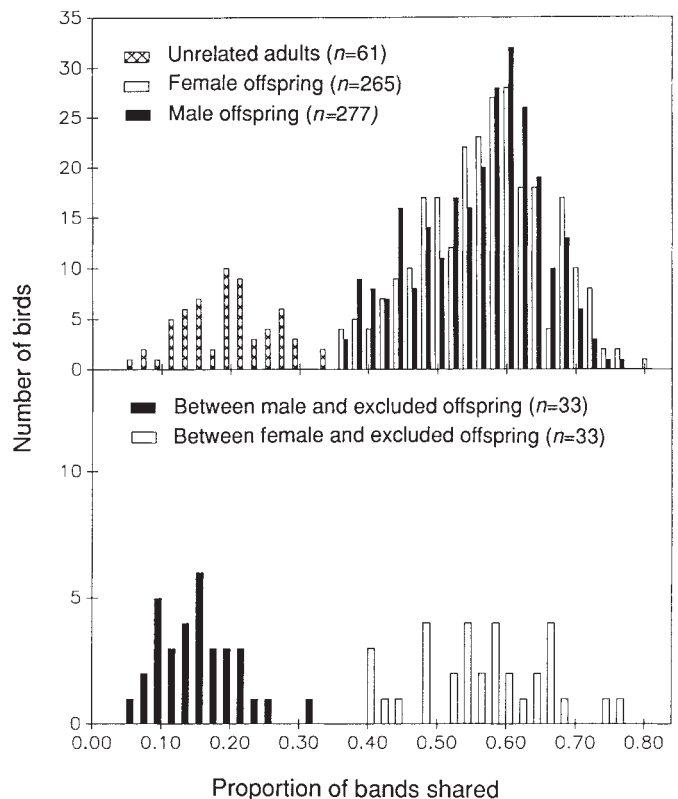
Mating system	Monogamy (1♂ 1♀)	Polygyny (1♂ 2♀)	
		Primary ♀	Secondary ♀
Mate guarding data	(n=28)	(n=8)	(n=9)
(a) % ♂ < 25 m	91 ± 1	79 ± 3	55 ± 5
(b) % ♂ initiated flights	37 ± 3	82 ± 2	87 ± 3
(c) % ♂ follows ♀	59 ± 4	53 ± 6	19 ± 3
Proportion of nests with extra-pair young	(n=25) 0.40	(n=6) 0.17	(n=5) 0.0

Mate guarding data (all during female's fertile period) include proportion of time male spent within 25 m of female (a), proportion of long flights (>10 m) initiated by male (b) and proportion of long flights initiated by female that are followed within 5 s by male (c). Data shown are mean ± s.e. (n = number of observed pairs). Pairs were watched continuously for several periods of 20–60 min during the breeding season (from start of nest building until female started incubation). At this time (early March to end of April) the leaves of the trees (mainly oak, birch and beech) were not yet developed, which made observations easier. Each pair was observed daily in 1990, every 2–3 days in 1991. The following observations were made: distance between male and female, flight initiation and following, copulations and copulation attempts, intrusions of other birds and the behaviour of these intruders. The fertile period of a female blue tit is defined as the period starting 5 days before the first egg is laid until the day the penultimate egg is laid¹⁷.

and there is a significant inverse correlation between the number of female intrusions in a male's territory and the number of extrusions of his own female (Fig. 2). Thus, females of 'attractive' males (with many intrusions from neighbouring females) do not leave their male when fertile, whereas females of 'unattractive' males (with no female intrusions) frequently visit the neighbours when fertile. If these visits lead to EPCs and finally to EPY, attractive males should suffer less lost paternity than unattractive males. To test this, we used those nests with both behavioural

FIG. 1 Frequency distributions of band sharing proportions between unrelated adults and between male or female and non-excluded offspring (a) and between male or female and excluded offspring (b). Band sharing proportion between male and excluded offspring is similar to that between unrelated adults, whereas proportion of bands shared with female was similar to that between non-excluded offspring and female. These young were then considered as EPY. In our sample there is no evidence for brood parasitism, because maternity is confirmed for all nestlings.

METHODS. Blood samples and dead young were stored at -70°C . For DNA extraction 5 μl of whole blood was added to 600 μl SET-buffer (0.15 M NaCl, 0.05 M Tris, 1 mM EDTA, pH 8.0) and incubated overnight with 15 μl proteinase K (10 mg ml^{-1}) and 10 μl 20% SDS at 55°C . The DNA was purified by phenol-chloroform extraction and recovered by ethanol precipitation at -20°C for 30 min. The DNA was collected by centrifugation, washed with 70% ethanol, vacuum dried and dissolved in 50 μl TE-buffer (10 mM Tris, 1 mM EDTA, pH 8.0) at 55°C . For each fingerprint about 2 μg DNA was digested overnight in the presence of 10 U *HinfI* enzyme (BRL). The DNA was digested to completion with the restriction enzyme *HinfI*, separated by agarose gel electrophoresis, transferred to a nylon membrane and hybridized with the radiolabelled minisatellite probe 33.15 (for more details on methods, see ref. 18). For parentage assignment we mainly followed the procedure described by Westneat¹². Mean number of scorable bands for the 33.15 probe was 20 (s.d. = 6.1, $n = 386$ individuals). If novel fragments were numerous (4–14) and the number of paternal (or maternal) specific bands was low (0–3), we considered this as an excluded offspring. If only one or two novel fragments were present and all other bands could be assigned to one or both parents, novel fragments were considered resulting from a mutation (mutation rate: 0.2% per fragment per meiotic event).



data and DNA-fingerprint results (data from 1990). The males suffering from lost paternity (EPY nests), received significantly fewer female visits than the males with no extra-pair young (NO-EPY nests) (mean female intrusions per hour: in EPY nests, 0.5 ± 0.3 s.d. ($n = 4$); NO-EPY nests, 2.0 ± 0.7 ($n = 6$), U -test, $U = 23.5$, $P < 0.025$), whereas the mean male intrusion rate is not significantly different (EPY nests, 1.7 ± 0.6 ($n = 4$); NO-EPY nests, 2.9 ± 1.7 ($n = 6$), $U = 16$, $P > 0.3$).

For 24 EPYs, paternity could be assigned using single-locus probes (Fig. 3), showing that (1) EPY are never reciprocal and (2) the standardized variance in male reproductive success was 69% higher when comparing realized to apparent success (Fig. 3).

Males that receive more female visits, recruit significantly more young (correlation between number of female intrusions per h into a male's territory and the number of recruits the following year from his nest: $r = 0.63$, $n = 8$, $P < 0.05$, one-tailed). This can be because they are better in providing paternal care, because they are genetically better, or because they occupy better territories. Males that lost paternity survived less well than males with no EPY in their nests (proportion surviving until next breeding season in EPY nests: 0.0, $n = 11$ and NO-EPY nests: 0.62, $n = 21$, Fisher's exact test, $P = 0.0006$), whereas there is no differential mortality for the females (EPY nests: 0.36, $n = 21$ and NO-EPY nests: 0.38, $n = 24$, Fisher's exact test, $P > 0.5$). A three-way loglinear test shows a significant interaction between survival, sex and EPY ($G = 8.42$, d.f. = 1, $P < 0.01$), suggesting that the difference in survival between males and females is not caused by differences in territory quality. Reduced recruitment and survival therefore both reflect male quality. Finally, males that lost paternity, but not their mates, had shorter tarsi than males that did not (mean tarsus length in mm: males from EPY nests, 15.1 ± 0.2 s.e. ($n = 11$); NO-EPY nests, 15.7 ± 0.2 ($n = 20$), $t = 2.19$, $P < 0.05$; females from EPY nests, 14.7 ± 0.2 ($n = 11$); NO-EPY nests, 14.8 ± 0.2 ($n = 20$), $t = 0.62$, $P > 0.5$). No differences in wing length were found (data not shown).

These results support the genetic quality hypothesis because males that lose paternity are more often left by their fertile female, recruit fewer young, survive less well and are smaller, whereas males that gain (extra-pair) paternity receive more female visits, recruit more young and survive better. A question that is left unanswered is how females are able to judge the quality of their mate and that of the neighbours. In one case a male injured one wing just before egg-laying. His female visited both neighbours and both shared paternity with the territorial male. After the young hatched, the territorial male died. Of the five males that died between 1 and 3 weeks after the female

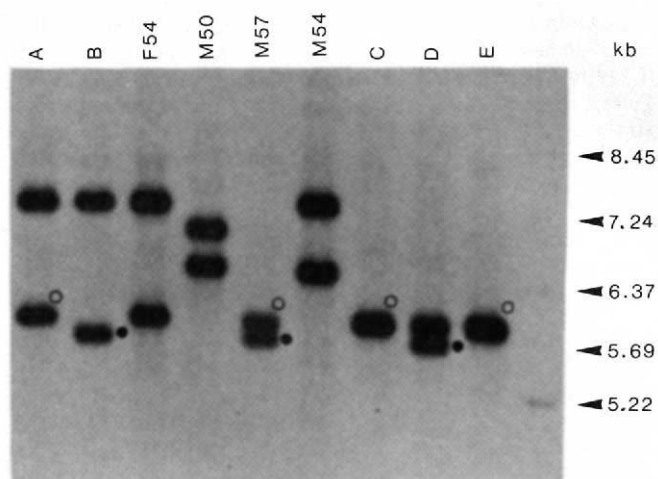


FIG. 3 Illustration of the use of the hypervariable single locus probe (SLP) cPcaMS14. The EPY (A-E) of nest 54 are shown with their mother (F54), her mate (M54) and two potential biological fathers (neighbours M50 and M57). Maternity was confirmed with probe 33.15 for all nestlings (band sharing proportions: 0.44–0.73, see Fig. 1). M57 has the 2 alleles found in the EPY (○ and ●), and is probably the biological father of the young. This result is confirmed with two other SLPs and with the 33.15 probe. Using SLPs, we identified 8 biological fathers of 24 EPY in 8 nests (probability of incorrect assignment between 1.1×10^{-4} and 1×10^{-5} ; our unpublished results). Two of them were attractive polygynous males (high number of female visits) with 17 and 20 young of their own (no EPY), respectively. Another male (no behavioural data) with seven legal young, fathered another nine EPY in three different nests. As a consequence, the standardized variance in male reproductive success in our population ($\text{variance}/(\text{mean})^2$; ref. 10) was 69% higher when based on realized success (0.27) as compared to apparent success (0.16).

METHODS. SLPs for blue tits were isolated (cPcaMS1 to cPcaMS14; unpublished data) using a method described elsewhere^{19,20}. SLPs were ³²P-radiolabelled using a random prime method²¹ and were hybridized at high stringency to the same membranes used with the 33.15 probe (see Fig. 1).

started incubation, four had a high proportion of EPY in their nest (5/6, 2/10, 3/3, 4/7). Perhaps females use the condition of a male during the breeding season as a clue to indicate his quality. □

Received 17 January; accepted 8 May 1992.

1. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B. G.) 136–179 (Aldine, Chicago, 1972).
2. Birkhead, T. R. & Møller, A. P. *Sperm Competition in Birds: Evolutionary Causes and Consequences* (Academic, London, 1992).
3. Smith, S. M. *Behaviour* **107**, 15–23 (1988).
4. Birkhead, T. R. et al. *Behav. Ecol. Sociobiol.* **27**, 315–324 (1990).
5. Møller, A. P. *Behav. Ecol. Sociobiol.* **27**, 23–29 (1990).
6. Westneat, D. F. et al. *Curr. Ornithol.* **7**, 331–369 (1990).
7. Bollinger, E. K. & Gavin, T. A. *Behav. Ecol. Sociobiol.* **29**, 1–7 (1991).
8. Møller, A. P. *Nature* **332**, 640–642 (1988).
9. Smith, H. G. & Montgomerie, R. *Behav. Ecol. Sociobiol.* **28**, 195–201 (1991).
10. Gibbs, H. L. et al. *Science* **250**, 1394–1397 (1990).
11. Sherman, P. W. & Morton, M. L. *Behav. Ecol. Sociobiol.* **22**, 413–420 (1988).
12. Westneat, D. F. *Behav. Ecol. Sociobiol.* **27**, 67–76 (1990).
13. Dhondt, A. A. in *Lifetime Reproduction in Birds* (ed. Newton, I. A.) 15–33 (Academic, London, 1989).
14. Dhondt, A. A. *Ibis* **129**, 327–334 (1987).
15. Birkhead, T. R. et al. *Nature* **334**, 60–62 (1988).
16. McCulloch, P. & Nelder, J. A. *Generalized Linear Models* (Chapman and Hall, New York, 1983).
17. Møller, A. P. *Behaviour* **100**, 92–104 (1987).
18. Berneman, Z. N. et al. *FEBS Lett.* **255**, 226–230 (1989).
19. Armour, J. A. L. et al. *Genomics* **8**, 501–512 (1990).
20. Bruford, M. W. et al. in *Molecular Genetic Analysis of Populations: A Practical Approach* (ed. Hoebul, A. R.) 225–269 (IRL, Oxford, 1992).
21. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).

ACKNOWLEDGEMENTS. We thank L. Bijnens for statistical help, A. J. Jeffreys for providing the minisatellite probe, I. van Pijlen for help with laboratory work, T. R. Birkhead for comments on the manuscript, E. Matthysen for discussion, and the Bracht family for allowing us to work on their estate. B.K. is Research Assistant of the Belgian National Fund for Scientific Research, G.R.V. holds an Instituut voor Wetenschappelijk Onderzoek in Nijverheid en Landbouw grant, T.B. is a Royal Society University Research Fellow, and C.V.B. is Research Associate of the Belgian National Fund for Scientific Research. This study was supported by a grant from the Belgian National Foundation for Scientific Research to A.D.

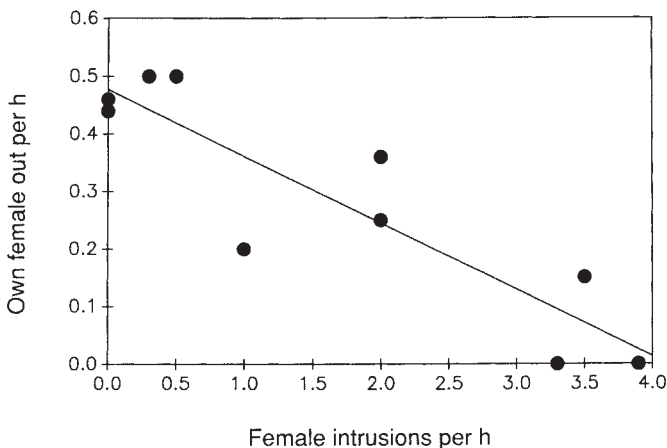


FIG. 2 Number of female intrusions per h of observation time (during their fertile period) into a male's territory versus the hourly visiting rate of his own female to neighbouring territories ($r^2 = 0.81$, $P < 0.001$, $n = 10$ males). Data from 1991.

Segregation of global and local motion processing in primate middle temporal visual area

Richard T. Born & Roger B. H. Tootell

Department of Neurobiology, Harvard Medical School,
220 Longwood Avenue, Boston, Massachusetts 02115, USA

THE early stages of primate visual processing appear to be divided up into several component parts so that, for example, colour, form and motion are analysed by anatomically distinct streams¹⁻³. We have found that further subspecialization occurs within the motion processing stream. Neurons representing two different kinds of information about visual motion are segregated in columnar fashion within the middle temporal area of the owl monkey. These columns can be distinguished by labelling with 2-deoxyglucose in response to large-field random-dot patterns. Neurons in lightly labelled interbands have receptive fields with antagonistic surrounds: the response to a centrally placed moving stimulus is suppressed by motion in the surround. Neurons in more densely labelled bands have surrounds that reinforce the centre response so that they integrate motion cues over large areas of the visual field. Interband cells carry information about local motion contrast that may be used to detect motion boundaries or to indicate retinal slip during visual tracking. Band cells encode information about global motion that might be useful for orienting the animal in its environment.

Our first hint that the middle temporal visual area (MT) had a modular organization came from studies with 2-deoxyglucose in which a large-field random-dot stimulus moving coherently at systematically varied directions, speeds and binocular disparities failed to produce uniform isotope uptake within this area (Fig. 1; ref. 4). We consistently obtained columnar patterns that were often band-like and regular but were occasionally more patchy, as in Fig. 1. We refer to the columns of high 2-deoxyglucose uptake as 'bands' and the intervening low-uptake regions as 'interbands'. This topography corresponds only roughly to the previously described pattern of cytochrome oxidase staining in owl monkey area MT⁵. We have chosen to use 2-deoxyglucose as our reference because it more directly reflects the functional properties of neurons and the cytochrome oxidase patterns are not consistently present in different primate species, whereas 2-deoxyglucose patterns are similar and consistent in both New and Old World monkeys.

The findings using 2-deoxyglucose prompted us to study the properties of single units in area MT with respect to their position in the band-interband architecture. In nine tangential and nine perpendicular microelectrode penetrations through area MT in anaesthetized, paralysed *Aotus* monkeys we recorded from 245 single units. The recording technique and stimulus generation

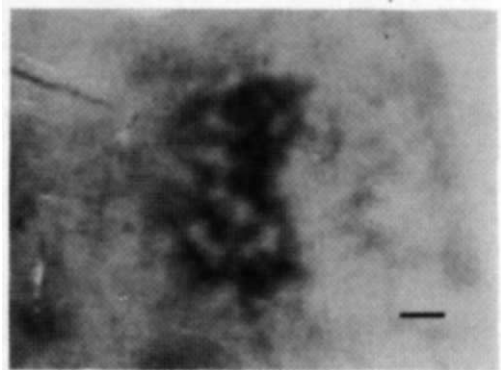
have been described⁶. After qualitative characterization of the best direction, speed and binocular disparity for each cell, the extent and nature of centre-surround interactions were measured using two sets of stimuli: (1) circular windows of varying diameter that contained random dots moving coherently in the cell's preferred direction (area summation test), and (2) a centre bar moving in the cell's preferred direction while the direction of motion of a surround patch of random dots was varied (surround direction test).

We found two major classes of neurons in area MT (Fig. 2e). The first class responded more vigorously as the stimulus size was increased (Fig. 2a). The excitatory effect of increased patch size often extended well beyond the borders of the receptive field as mapped with a single bar, thus indicating the presence of a surround whose action reinforced the centre's directional response. Small bars were usually ineffective stimuli for these cells (Fig. 2b, bar); they responded best in the surround direction test when the large surround field of random dots moved in the same direction as that preferred by the centre (Fig. 2b). The second major class failed to respond to large fields of coherently moving stimuli, indicating the presence of a surround that was antagonistic to the centre response (Fig. 2c). The suppressive surround was sometimes direction-selective, so that motion in the surround at the same direction and speed as the centre motion inhibited the response, whereas surround motion in the opposite direction actually enhanced the response over that to the centre stimulus alone (Fig. 2d). In other cells, any motion in the surround was inhibitory regardless of its direction (Fig. 4b, top, right-hand graph).

During the nine tangential penetrations, we consistently found that neurons with similar centre-surround interactions were clustered (Fig. 3). In three of these penetrations we used 2-deoxyglucose to label the band-interband architecture after completing the single unit recording. When the electrode track was reconstructed and superimposed on an autoradiograph of the 2-deoxyglucose-labelled tissue, the cells whose receptive fields had antagonistic surrounds were located in the light interbands, and the cells whose receptive fields had reinforcing surrounds were located predominantly in the dark bands (Fig. 3a). By staining the sections for cytochrome oxidase⁷ or cresyl violet or both, we determined the laminar position of each unit recorded (Fig. 3b); this allowed us to tell whether changes in unit properties were due to horizontal or laminar position.

The columnar nature of the band-interband architecture was confirmed by electrode penetrations perpendicular to the cortical surface. In these penetrations we always found that one receptive field type dominated, with virtually all the exceptions in the upper parts of layers 4 and 6 (Fig. 4). We determined the laminar position of 118 neurons during the nine perpendicular penetrations. Of these, only 19 neurons (16%) were discordant with their column type (almost always a neuron lacking an antagonistic surround, in an interband), and 16 of these 19

FIG. 1 Band-interband architecture of area MT as shown by the 2-deoxyglucose method. An anaesthetized, paralysed *Aotus* monkey was injected intravenously with ¹⁴C-labelled-deoxyglucose while it binocularly viewed a large field (60° in diameter) of random dots that were systematically drifted in all directions and at a variety of speeds. The 2-deoxyglucose protocol and subsequent histological processing have been described²². Scale bar, 1 mm.



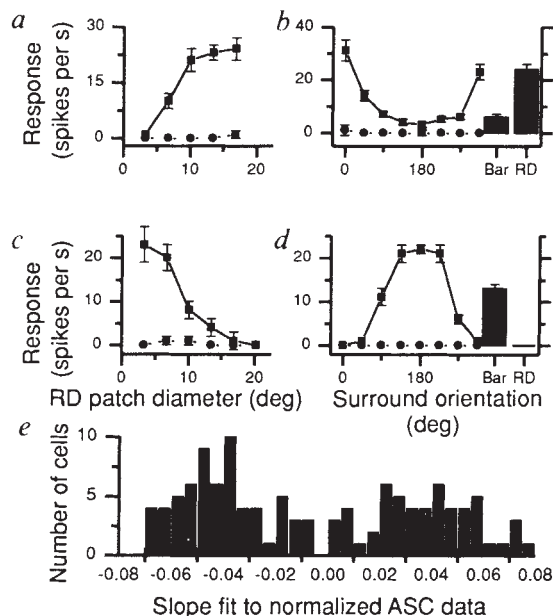


FIG. 2 Two major cell classes in area MT. *a–d*, Stimulus-response curves from one band cell (*a, b*) and one interband cell (*c, d*). *a* and *c*, Area summation curves showing each cell's response to patches of coherently moving random dots. Increasing values on the abscissa indicate progressively larger circular areas of the visual field stimulated with random dots (dot density was constant). The smallest patch size was always well within the borders of the cell's receptive field as mapped by a single bar. *b* and *d*, Surround direction tuning curves in which the centre of each cell's receptive field was stimulated with an optimally moving bar while the direction of motion of a surrounding field of random dots was varied. Zero degrees means that the surround dots moved in the same direction as the centre bar, and 180 degrees means that the surround dots moved in the opposite direction. The centre-surround interactions shown in the line plots should be compared with the adjacent column plots which show the cell's response to a centre bar moving on a stationary field of random dots (Bar) and a large field ($>20^\circ$) of random dots moving in the cell's preferred direction while the centre bar was stationary (RD). Note that in *d*, the presence of surround dots moving in a direction opposite to the best direction for the centre bar increased the response over that to the bar alone. The filled squares connected by solid lines indicate the response while the stimulus was on and moving, and the filled circles connected by dotted lines indicate the cell's level of spontaneous activity just prior to each stimulus presentation. Each point is the average of five trials. Error bars are $\pm$ one standard deviation. *e*, Histogram of the slopes of regression lines fit to the normalized area summation curves for 118 unselected single units. It is extremely unlikely that this sample was drawn from a single normally distributed population ($P < 0.00001$, Shapiro-Wilk W test for normality). Quantile analysis gave a sigmoidal curve with the inflection point at zero, indicating a bimodal distribution.

(84%) were in layer 4 or the top of layer 6. In layers 2 and 3, only one of 48 (2%) units failed to match its column type.

Neurons whose receptive fields manifest centre-surround interactions for moving stimuli have been described previously in primate area MT^{8–11}, in cat superior colliculus¹² and in pigeon optic tectum^{13,14}. In primate area MT there is a tendency for antagonistic surround inhibition to distribute laminarly, being more prevalent in the supra- and infragranular layers and tending to avoid layer 4 (refs 10, 11). Our results in the owl monkey

agree with these results, and in addition we find that cells in the uppermost part of layer six also often lack antagonistic surrounds, even in the interbands. Both of these thin layers receive inputs from striate cortex¹⁵.

In addition to a laminar organization, we found a horizontal clustering of centre-surround interactions, which was constant with respect to cortical depth. The overall architecture is thus predominantly columnar with thin gaps in the striate recipient layers. This situation is analogous to the columnar orientation

FIG. 3 Reconstruction of a tangential penetration through area MT. *a*, Electrode track superimposed on the 2-deoxyglucose-labelled band-interband pattern. Each single unit is represented by a dot whose brightness indicates the degree of surround inhibition. The degree of surround inhibition was measured as the slope of the regression line fit to the normalized area summation curve, as in Fig. 2*e*. If the slope was greater than or equal to zero, the cell was classified as non-surround inhibited (white); if the slope was less than zero but greater than -0.03 , the cell was deemed weakly surround inhibited (grey); and if the slope was less than or equal to -0.03 (indicating at least a halving of the cell's best response), the cell was called strongly surround-inhibited (black). *b*, The same penetration at an expanded scale showing the laminar position of each unit (top). The solid dark bars represent the result of a densitometric scan along the electrode track taken from the 2-deoxyglucose image in *a*. Circles show the degree of surround inhibition (classified as described in *a*) for each unit. Note that the cluster of three interband cells that lack surround inhibition (just after $3,000\ \mu\text{m}$) was located in the uppermost part of layer 6, which receives inputs from striate cortex¹⁵. Abbreviations: A, anterior; P, posterior; D, dorsal; V, ventral; STS, superior temporal sulcus; SI, surround inhibition.

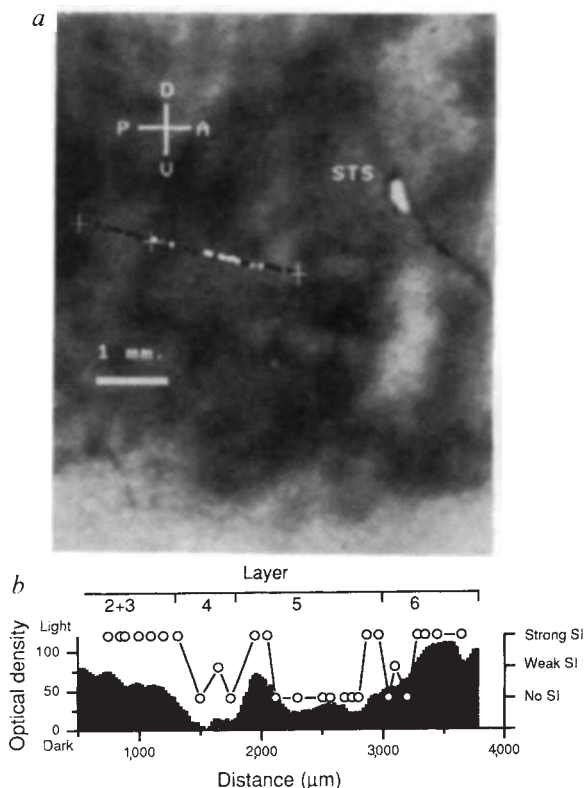
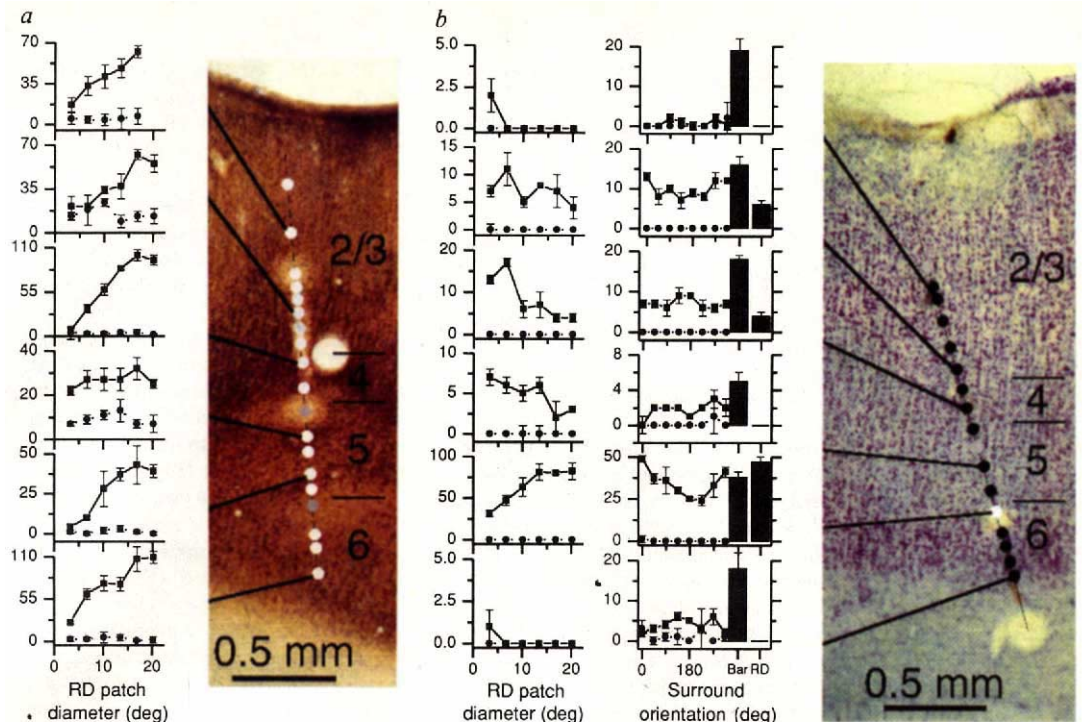


FIG. 4 Electrode track reconstructions of two vertical penetrations through area MT showing the columnar nature of the band-interband architecture. *a*, Cresyl violet/cytochrome oxidase histology of a presumptive band penetration with single units coded for the strength of surround inhibition as in Fig. 2. Both lesions are visible in this section: one near the middle of layers 2/3 and the other between layers 4 and 5. Immediately to the left are area summation curves of six single units recorded at the positions indicated. The stimulus protocol used to generate these curves was identical to that for Fig. 1*a* and *c*. *b*, Corresponding data for a presumptive interband penetration superimposed on the cresyl violet histology. Two of the three lesions are clearly visible in this section: one between layers 5 and 6 and another, larger lesion in the white matter. A third lesion, faintly visible in this section near the top of layer 4, is more obvious in the adjacent section (not shown). For this penetration, surround direction tuning curves are also shown because some of the surround-inhibited units gave poor responses even to the smallest diameter random dot fields (for example, topmost unit graphed.).



The presence of a strong non-direction-selective surround inhibition can be seen by comparing the cell's response to a centre bar alone (Bar) with the responses when there is surround motion. Any motion in the surround, regardless of direction, was inhibitory.

architecture in macaque striate cortex where parvocellular geniculate input layers can lack orientation selectivity¹⁶. This does not mean that MT neurons are strictly dichotomous with respect to surround antagonism. Indeed, the histogram of area summation curve slopes (Fig. 2*e*) shows cells with intermediate degrees of surround antagonism, consistent with the fact that band-interband boundaries are not perfectly sharp (Fig. 3*b*).

We believe that this columnar organization indicates that parallel processing is taking place within the motion pathway. The neurons with antagonistic surrounds in area MT seem to be calculating local motion contrast, analogous to the calculation of local intensity contrast performed by retinal ganglion cells¹⁷. This kind of information could be useful for perceptual popout¹⁸, as a signal from an object boundary moving with respect to the background⁹, due either to true object motion or to observer motion when the object and background are at different depths (motion parallax). Another possibility is that local motion contrast cells whose receptive fields are near the fovea provide, during visual tracking, an image velocity signal, which behavioral experiments have indicated is a necessary input to the smooth pursuit system¹⁹.

Cells within the bands differ from interband cells in that they sum motion cues over relatively large regions of the visual field. This type of operation is necessary for the calculation of global motion properties, such as optic flow, which, combined with vestibular and proprioceptive information, help the animal orient to and navigate within its environment. It is also clear that cells with such large excitatory integrating receptive fields could provide the inhibitory surrounds for interband cells, although we have no direct evidence that band cells serve this role.

This anatomical segregation within MT may be the precursor to a more complete parcellation of these two broad categories of information about visual motion. The principal projection area of MT in the macaque, area MST, consists of two separate representations of the visual field, MSTd and MSTl, whose receptive field properties resemble those of band and interband cells, respectively²⁰. It has been recently reported that a dorsal subdivision of area FST (the owl monkey homologue of MST) receives inputs from an array of cell clusters within area MT (ref. 21). It would be interesting to know if these parallel projections map on to the band-interband architecture. □

Received 9 March; accepted 30 April 1992.

- Livingstone, M. S. & Hubel, D. H. *Science* **240**, 740-749 (1988).
- DeYoe, E. A. & Van Essen, D. C. *Trends Neurosci.* **11**, 219-226 (1988).
- Maunsell, J. H. R. & Newsome, W. T. *A. Rev. Neurosci.* **10**, 363-440 (1987).
- Tootell, R. B. H. & Born, R. T. *Invest. ophthalmol. visual Sci.* **31**, 238 (1990).
- Tootell, R. B. H., Hamilton, S. L. & Silverman, M. S. *J. Neurosci.* **5**, 2786-2800 (1985).
- Born, R. T. & Tootell, R. B. H. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7066-7070 (1991).
- Wong-Riley, M. T. T. *Brain Res.* **171**, 11-28 (1979).
- Baker, J. F., Petersen, S. E., Newsome, W. T. & Allman, J. M. *J. Neurophysiol.* **45**, 397-416 (1981).
- Allman, J., Miezin, F. & McGuinness, E. *Perception* **14**, 105-126 (1985).
- Tanaka, K., Hikosaka, K. & Saito, H. *J. Neurosci.* **6**, 134-144 (1986).
- Lagae, L., Gulyas, B., Raiguel, S. & Orban, G. A. *Brain Res.* **496**, 361-367 (1989).
- Sterling, P. & Wickelgren, B. G. *J. Neurophysiol.* **32**, 1-15 (1969).

- Frost, B. J., Scille, P. L. & Wong, S. C. P. *Expl Brain Res.* **43**, 173-185 (1981).
- Frost, B. J. & Nakayama, K. *Science* **220**, 744-745 (1983).
- Rockland, K. S. *Vis. Neurosci.* **3**, 155-170 (1989).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215-243 (1968).
- Kuffler, S. W. *J. Neurophysiol.* **16**, 37-68 (1953).
- Treisman, A. *Sci. Am.* **255**, 106-115 (1986).
- Krauzlis, R. J. & Lisberger, S. G. *Neural Comp.* **1**, 116-122 (1989).
- Wurtz, R. H., Duffy, C. J. & Roy, J. P. *Cold Spring Harb Symp. quant. Biol.* **LV**, 45 (1990).
- Kaas, J. H. & Morel, A. E. *Invest. ophthalmol. vis. Sci.* **33**, 1219 (1992).
- Tootell, R. B. H., Hamilton, S. L., Silverman, M. S. & Switkes, E. *J. Neurosci.* **8**, 1500-1530 (1988).

ACKNOWLEDGEMENTS. We thank G. Blasdel, A. Das, D. Hubel, M. Livingstone and P. Ma for discussion and comments on the manuscript. R. Huber provided assistance with statistical analysis.

Targeting of cell-surface β -amyloid precursor protein to lysosomes: alternative processing into amyloid-bearing fragments

Christian Haass, Edward H. Koo, Angela Mellon, Albert Y. Hung & Dennis J. Selkoe

Departments of Neurology and Pathology and Program in Neuroscience, Harvard Medical School, and Center for Neurologic Diseases, Division of Neurology, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

PROGRESSIVE cerebral deposition of the amyloid β -peptide is an early and invariant feature of Alzheimer's disease. The β -peptide is released by proteolytic cleavages from the β -amyloid precursor protein (β APP)¹, a membrane-spanning glycoprotein expressed in most mammalian cells. Normal secretion of β APP involves a cleavage in the β -peptide region^{2,3}, releasing the soluble extramembranous portion^{4,5} and retaining a 10K C-terminal fragment in the membrane⁶. Because this secretory pathway precludes β -amyloid formation, we searched for an alternative proteolytic processing pathway that can generate β -peptide-bearing fragments from full-length β APP. Incubation of living human endothelial cells with a β APP antibody revealed internalization of mature β APP from the cell surface and its targeting to endosomes/lysosomes. After cell-surface biotinylation, full-length biotinylated β APP was recovered inside the cells. Purification of lysosomes directly demonstrated the presence of mature β APP and an extensive array of β -peptide-containing proteolytic products. Our results define a second processing pathway for β APP and suggest that it may be responsible for generating amyloid-bearing fragments in Alzheimer's disease.

The presence of a consensus sequence known to mediate clathrin coated pit internalization in the β APP carboxyl terminus (Asn-Pro-Thr-Tyr, amino acids 759-762) indicates that β APP could be reinternalized from the cell surface like a variety of receptors⁷. To investigate this possibility, we used human umbilical vein endothelial cells (HUVEC (ref. 8)), which have the advantage of being primary human cells that express large amounts of β APP. RNase protection assays (Fig. 1b) showed

that HUVEC transcribe the 751- and 770-amino-acid isoforms (β APP 751 and β APP 770) of β APP, which are derived from alternative splicing of the β APP pre-messenger RNA⁹⁻¹¹. Comparison of the amounts of β APP in HUVEC and stably transfected human kidney 293 cells^{5,6} by immunoblotting showed that HUVEC synthesize substantial amounts of β APP (Fig. 1c). Fluorescence immunocytochemistry using the C-terminal antibody, C7 (ref. 12), and the affinity-purified mid-region antibody, B5 (ref. 5), demonstrated strong perinuclear staining representing primarily an endoplasmic reticulum and Golgi-like localization (Fig. 2a, c). Cell identity was confirmed by double labelling with a monoclonal antibody against von Willebrand factor, which strongly stained Weibel-Palade bodies¹³, specialized Golgi vesicles characteristic of HUVEC (Fig. 2b, d).

To reduce the breakdown of any β APP that might be reinternalized and targeted to lysosomes, we treated the cells with the lysosomal protease inhibitor, leupeptin. This produced a coarse vesicular staining for β APP (Fig. 2e, g), indicating a possible localization to late endosomes/lysosomes. The formation of von Willebrand factor-positive Weibel-Palade bodies was not altered by leupeptin, indicating that Golgi functions and protein traffic were normal (Fig. 2f, h). This conclusion was further supported by pulse-chase experiments, which showed no difference in the maturation of full-length β APP in HUVEC cultured with or without leupeptin (data not shown). Moreover, leupeptin does not inhibit the secretion of soluble β APP (ref. 14, and C.H. *et al.*, data not shown).

To determine whether this apparent lysosomal localization of β APP was due to reinternalization, we added affinity-purified antibody B5 to living cells, expecting it to bind to the extramembranous domain (Fig. 1a) and be reinternalized as a β APP/antibody complex which could be targeted to lysosomes. This experiment (Fig. 2i) produced lysosome-like staining identical to that already described (Fig. 2e, g). The late endosomal/lysosomal localization of internalized β APP was confirmed by simultaneous internalization of Texas red-conjugated albumin (Fig. 2j), which is known to be targeted to late endosomes/lysosomes by fluid phase endocytosis¹⁵. Cells incubated with antibody C7, which recognizes only the intracellular domain (Fig. 1a), did not stain, ruling out nonspecific antibody uptake (data not shown).

To confirm that the B5 antibodies did not redirect β APP out of its normal pathway, we directly showed reinternalization by labelling cell-surface proteins with biotin (Fig. 3). Following

FIG. 1 Analysis of β APP expression in HUVEC. a Schematic structure of β APP 770. Vertical lines indicate the transmembrane domain; white boxes designate the Kunitz protease inhibitor (KPI) domain and adjacent exon; the black box represents the signal sequence and the hatched box the β -peptide. Horizontal bars indicate the antigens for antibodies B5 (amino acids 519-667), C7 (amino acids 751-770) and 22C11 (Boehringer Mannheim). Vertical arrow indicates site of constitutive cleavage that results in β APP secretion. **b** RNase protection assay shows that HUVEC (lane 1) contain β APP 770 and β APP 751 mRNAs in similar amounts. RNA derived from human kidney 293 cells stably transfected with β APP 751 (lane 2), β APP 695 (lane 3) or untransfected (lane 4) were used as standards. **c** Immunoblot analysis¹² of β APP expressed by HUVEC (lane 1), 293 cells transfected with β APP 751 (lane 2) or β APP 695 (lane 3) or untransfected (lane 4), using antibody C7. A total protein extract (100 μ g) was electrophoresed on a 7% polyacrylamide gel and transferred to nitrocellulose.

METHODS. The complementary DNA construct for RNase protection was generated by polymerase chain reaction (PCR) using primers corresponding to base pairs (bp) 640-660 and 1,200-1,219 of β APP 770 (ref. 12). The PCR-amplified fragment was subcloned into the pGEM3 vector using the primer-derived *Hind*III and *Bam*HI restriction sites. For RNase protection assays, antisense RNA was transcribed *in vitro* from the *Bam*HI-linearized construct using a SP6 polymerase kit (Boehringer Mannheim).

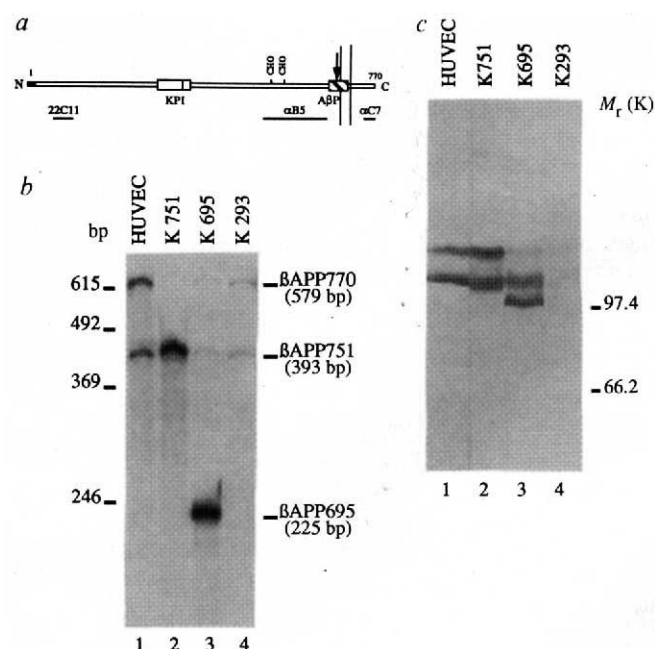
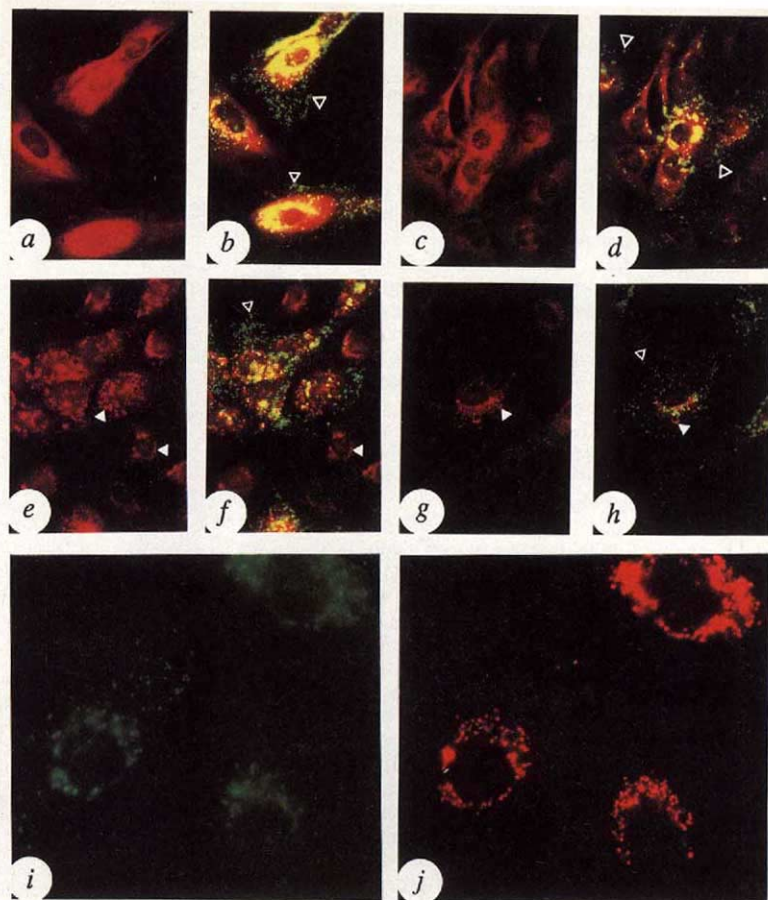


FIG. 2 Immunocytochemical detection of β APP in lysosomes of HUVEC. HUVEC were double-stained with a monoclonal antibody against von Willebrand factor (green) and with C7 antibody (red). *a*, β APP staining; *b*, double exposure to visualize both von Willebrand factor in Weibel–Palade bodies (triangles) and β APP; *c* and *d*, same as *a* and *b*, but using affinity-purified B5 antibody; *e* and *f*, same as *a* and *b* but using HUVEC treated for 12 h with $20 \mu\text{g ml}^{-1}$ leupeptin. Strong vesicular C-terminal reactivity (filled triangles) is observed in late endosomes/lysosomes (confirmed by internalization of Texas red-conjugated albumin in *j*, *g* and *h*). Same as *e* and *f*, but using B5 antibody after a 12-h leupeptin treatment ($20 \mu\text{g ml}^{-1}$). B5 (filled triangles) gives the same late endosomal/lysosomal staining as C7, ruling out the possibility that only the 10K C-terminal fragment rather than full-length β APP is targeted to lysosomes. Note that leupeptin has no influence on the generation of the von Willebrand factor-positive Weibel–Palade bodies (*f*, *h*). Magnification in *a*–*h*, $\times 830$. *i*, Living cells incubated for 12 h with B5 antibody ($1:300$) in the presence of $20 \mu\text{g ml}^{-1}$ leupeptin. Complexes of B5 bound to cell-surface β APP and reinternalized were detected in late endosomes/lysosomes by staining with a fluorescein isothiocyanate-conjugated secondary antibody alone (green). *j*, The late endosomal/lysosomal localization of β APP was confirmed by the simultaneous internalization of Texas red-conjugated albumin (red; same cells as in *i*). The two molecules colocalize extensively but not completely, as expected from two different mechanisms of internalization (coated pit-mediated reinternalization versus fluid-phase endocytosis). Magnification in *i* and *j*, $\times 2,350$.

METHODS. Leupeptin (Sigma) was added from a 10 mg ml^{-1} stock solution in medium to cultured cells for 12 h. For studies on the reinternalization of the antigen–antibody complex²³ and for immunocytochemical detection, HUVEC were grown on collagen-coated coverslips. Immunocytochemistry has been described¹².



biotinylation, cell lysates and medium were immunoprecipitated with either C7 antibody (lysates) or B5 antibody (media) (Fig. 3). Biotinylated soluble β APP was present in the medium (Fig. 3, lanes 4 and 5), directly demonstrating that cleavage of cell-surface β APP can give rise to the secreted form of the molecule. Biotinylated mature β APP was recovered in the lysates (Fig. 3, lanes 2 and 3), clearly showing reinternalization of the uncleaved molecule in the presence or absence of leupeptin. This result shows that normal movement of β APP to endosomes/lysosomes does not depend on effects of leupeptin.

To confirm directly the localization of internalized β APP, lysosomes were purified from both HUVEC and 293 cells stably transfected with β APP 751 (ref. 5) (Fig. 4). The identity of subcellular fractions was confirmed by detection of the lysosome-specific marker, cathepsin D (Fig. 4*a*). Full-length β APP and the constitutive C-terminal fragment of M_r 10,000 (10K) were detected in purified lysosomes in the absence of leupeptin treatment (Fig. 4*b* and *c*, lanes 3). Moreover, lysosomes isolated after cell-surface biotinylation contained mature full-length biotinylated β APP, confirming reinternalization and targeting to lysosomes (data not shown). Comparison of various subcellular fractions revealed enrichment of the 10K fragment within lysosomes (compare lanes 1 and 3 in Fig. 4*b* and *c*), showing that this non-amyloidogenic proteolytic product is, as expected, targeted to lysosomes for its final degradation. In lysosomes purified from leupeptin-treated cells, we observed a marked enrichment of full-length β APP, the 10K fragment, and a striking array of slightly larger C-terminal fragments which should contain the intact β -peptide (Fig. 4*b*–*d*). These results illustrate the stabilization of β APP as well as its C-terminal breakdown products by inhibiting lysosomal proteases. Besides the 10K fragment generated by secretory cleavage, the most prominent proteolytic fragments stabilized by leupeptin have M_s $\sim 22\text{K}$ (Fig. 4*b* and *c*, asterisk). It is of interest that a stable C-terminal

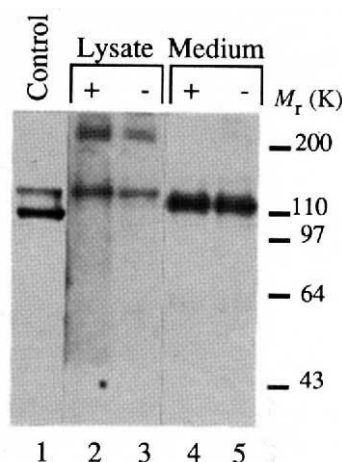
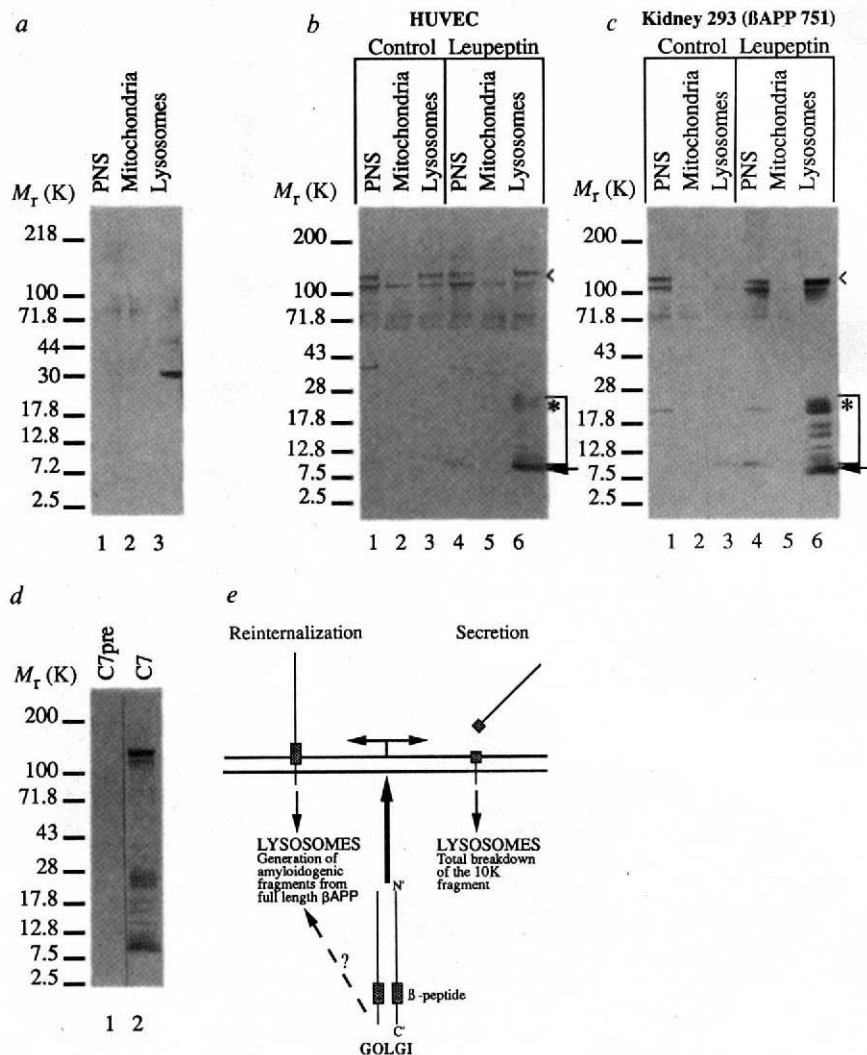


FIG. 3 Surface biotinylation of β APP in HUVEC followed by reinternalization or secretion. Biotinylated β APP species were immunoprecipitated with C7 antibody from total cell lysates treated either with (lane 2) or without (lane 3) leupeptin ($20 \mu\text{g ml}^{-1}$). Biotinylated soluble β APP was immunoprecipitated with B5 from the medium of cells treated with (lane 4) or without (lane 5) leupeptin ($20 \mu\text{g ml}^{-1}$). As a β APP standard, $50 \mu\text{g}$ of a total cell lysate of 293 cells transfected with β APP 751 (ref. 5) (lane 1) was run in parallel and immunoblotted with monoclonal antibody 22C11 (ref. 4) (Fig. 1*a*). METHODS. Living HUVEC were surface biotinylated according to ref. 24. After biotinylation, cells were incubated for 2 h at 37°C to allow internalization and then trypsinized to remove any biotinylated β APP remaining on the cell surface. To check that surface biotinylated β APP had all been removed, cells were trypsinized immediately after biotinylation (thereby preventing internalization) and washed, lysed and immunoprecipitated with C7; under these conditions virtually no β APP was recovered in the lysate (data not shown). The immunoprecipitated biotinylated β APP species were separated on a 7% polyacrylamide gel, transferred to nitrocellulose and detected autoradiographically using ^{125}I -labelled-streptavidin.

FIG. 4 Detection of β APP in purified lysosomes. **a**, Post-nuclear supernatant (PNS; lane 1), mitochondria (lane 2) and lysosomes (lane 3) (20 μ g of each) isolated from HUVEC were immunoblotted with a polyclonal antibody to cathepsin D. Note the strong enrichment of mature cathepsin D (CatD) and procathepsin D (proCatD) in the lysosomal fraction. **b**, Post-nuclear supernatant (PNS) (lanes 1, 4), mitochondria (lanes 2, 5) and lysosomes (lanes 3, 6) (20 μ g of each) isolated from HUVEC treated with 20 μ g ml⁻¹ leupeptin (lanes 4–6) or without leupeptin (lanes 1–3) were immunoblotted with C7 antibody. **c**, Post-nuclear supernatant (lanes 1, 4), mitochondria (lanes 2, 5) and lysosomes (lanes 3, 6) (20 μ g of each) isolated from β APP 751-transfected kidney 293 cells treated either with 20 μ g ml⁻¹ leupeptin (lanes 4–6) or without leupeptin (lanes 1–3) were immunoblotted with C7. Arrowhead, Full-length mature β APP; bracket, low M_r , potentially amyloidogenic fragments; asterisks, 22K fragments; arrow, constitutive non-amyloidogenic 10K fragment^{2,6}. Note that lysosomes purified from HUVEC (**b**) and transfected 293 cells (**c**) contain the same pattern of potentially amyloidogenic breakdown products (lanes 6). The band at ~21K in **c**, lanes 1 and 4, is nonspecific and not derived from β APP (data not shown); note its absence in HUVEC PNS (**b**, lanes 1, 4). **d**, Isolated lysosomes (20 μ g) from 293 cells transfected with β APP 751 were immunoblotted with antibody C7 with (lane 1) or without (lane 2) preabsorption with the C7 peptide immunogen, showing that C7 specifically recognizes full-length β APP, the potentially amyloidogenic breakdown products and the 10K fragment (lane 2). **e**, Model showing the two normal processing pathways for β APP directly demonstrated in this study: cleavage and secretion of soluble β APP at the cell surface, or reinternalization and lysosomal targeting of full-length β APP. Dashed line indicates a potential direct pathway of β APP from the Golgi to the lysosome. METHODS. Lysosomes were purified according to ref. 25. This method yields highly purified lysosomes (see ref. 26, for example). Mitochondria were isolated as a separate fraction from the same gradient as the lysosomes. All fractions were electrophoresed on 10–20% Tris/Tricine gels and transferred to nitrocellulose. Antibody preabsorption has been described¹².



β APP fragment of this size has been detected in purified human brain microvessels¹⁶. The bands between ~11K and ~15K (Fig. 4b–d) probably represent the β APP fragments described in human brain¹⁷ and in transfected 293 cells¹⁴. These fragments are also stabilized by leupeptin treatment¹⁴, indicating a possible lysosomal origin. We have shown that these bands are found in lysosomes but that they represent only a minority of potentially amyloidogenic fragments. This lysosomal targeting is consistent with earlier suggestions of a lysosomal localization of β APP (refs 18, 19).

Our observations define a second constitutive pathway of β APP processing and suggest the model shown in Fig. 4e. Full-length β APP molecules reach the cell surface after maturation, where they can either be cleaved to release soluble β APP into the medium or reinternalized to lysosomes, thus providing a substrate containing the intact β -peptide region to lysosomal proteases. We have shown directly that the latter can yield small β -peptide-bearing fragments; such peptides could potentially be released into the extracellular space^{20,21}. Our data do not rule out the existence of an additional pathway that targets β APP directly from the late *trans*-Golgi network to the lysosome.

The effect of β APP mutations in familial Alzheimer's disease²² could alter the balance of reinternalization versus secretion or

the stability of amyloidogenic fragments within lysosomes. Our data suggest that minor changes in the relative use or degradative capacity of these two normal pathways may result in the gradual accumulation of amyloidogenic peptides characteristic of Alzheimer's disease and normal brain ageing. □

Received 12 March; accepted 29 April 1992.

- Kang, J. et al. *Nature* **325**, 733–736 (1987).
- Esch, F. et al. *Science* **248**, 1122–1124 (1990).
- Sisodia, S. S., Koo, E. H., Beyreuther, K., Unterbeck, A. & Price, D. L. *Science* **248**, 492–495 (1990).
- Weidemann, A. et al. *Cell* **57**, 115–126 (1989).
- Oltschendorf, T. et al. *J. Biol. Chem.* **265**, 4492–4497 (1990).
- Selkoe, D. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7341–7345 (1988).
- Chen, W. J., Goldstein, J. S. & Brown, M. S. *J. Biol. Chem.* **265**, 3116–3123 (1990).
- Gimbrone, M. A., Cotran, R. S. & Folkman, J. *J. Cell Biol.* **60**, 673–684 (1974).
- Ponte, P. et al. *Nature* **331**, 525–527 (1988).
- Tanzi, R. E. et al. *Nature* **331**, 528–530 (1988).
- Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S. & Ito, H. *Nature* **331**, 530–532 (1988).
- Haass, C., Hung, A. Y. & Selkoe, D. J. *J. Neurosci.* **11**, 3783–3793 (1991).
- Weibel, E. B. & Palade, M. D. *J. Cell Biol.* **23**, 101–112 (1964).
- Golde, T. et al. *Science* **255**, 728–730 (1992).
- Parton, R. G. et al. *J. Cell Biol.* **113**, 261–274 (1991).
- Tamaoka, A., Kalaria, R. N., Lieberburg, I. & Selkoe, D. J. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1345–1349 (1992).
- Estus, S. et al. *Science* **255**, 726–728 (1992).
- Benowitz, L. I. et al. *Exp. Neurol.* **106**, 237–250 (1989).
- Cole, G. M., Huynh, T. V. & Saitoh, T. *Neurochem. Res.* **10**, 933–939 (1989).
- Buktenica, S., Olenick, S. J. & Frankfurter, A. *J. Biol. Chem.* **262**, 9469–9476 (1987).

21. Baskin, F., Rosenberg, R. G. & Greenberg, B. D. *J. Neurosci. Res.* **29**, 127–132 (1991).
22. Goate, A. *et al. Nature* **349**, 704–706 (1991).
23. Hunziker, W., Hartler, C., Matter, K. & Mellman, I. *Cell* **66**, 907–920 (1991).
24. Lisanti, M. P., Sargiacomo, M., Graeve, L., Saltiel, A. R. & Rodriguez-Boulant, E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9557–9561 (1988).
25. Storrie, B. & Madden, E. *Meth. Enzym.* **182**, 203–210 (1990).
26. Dice, J. F. & Terlecky, S. P. *Crit. Rev. Therap. Drug Carrier Systems* **7**, 211–233 (1990).

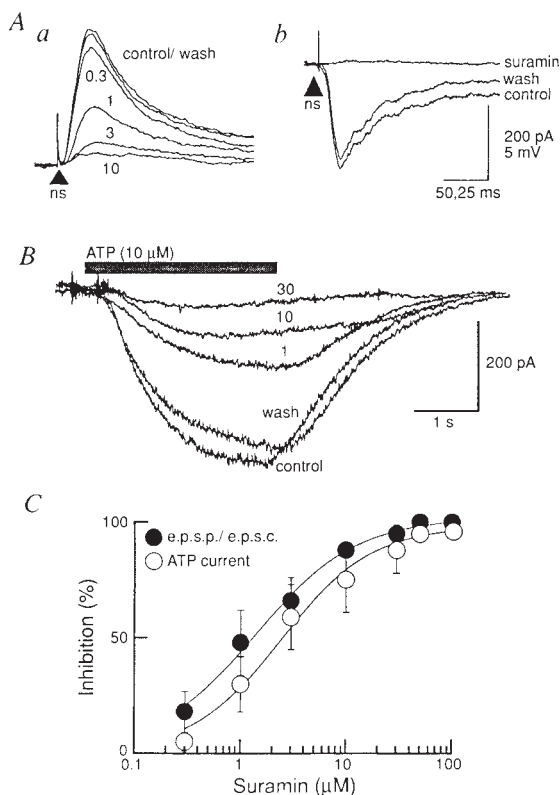
ACKNOWLEDGEMENTS. We thank K. W. Case and M. A. Gimbrone for tissue for isolation of HUVEC, P. Ward and I. Lieberburg for antibody B5, and J. F. Dice, T. Stossel and K. Kosik for suggestions. Supported by grants from the NIH and Athena Neurosciences (D.J.S.), the Alzheimer's Association (E.K.), the Deutsche Forschungsgemeinschaft/Boehringer Ingelheim Fonds (C.H.) and Merck Sharp and Dohme Research Laboratories (A.Y.H.).

ATP mediates fast synaptic transmission in mammalian neurons

Richard J. Evans, Victor Derkach
& Annmarie Surprenant

Vollum Institute, Oregon Health Sciences University, Portland,
Oregon 97201, USA

IN addition to its diverse functions inside cells, ATP can act at several types of cell-surface receptor^{1–3}. One of these (P_{2X}-purinoceptor) is believed to be a ligand-gated cation channel^{1–6}. The presence of P_{2X} receptors on autonomic, sensory and central neurons suggests that ATP might be released to act as a fast excitatory synaptic transmitter. Here we record excitatory synaptic potentials and currents from cultured coeliac ganglion neurons which are mimicked by ATP, blocked by the P₂-purinoceptor antagonist suramin, desensitized by α,β -methylene-ATP and unaffected by antagonists acting at nicotine, 5-hydroxytryptamine, N-methyl-D-aspartate (NMDA), non-NMDA glutamate, γ -aminobutyric acid (GABA), noradrenaline or adenosine receptors. We conclude that ATP is the neurotransmitter at this neuro-neuronal synapse.



Fast excitatory postsynaptic potentials (e.p.s.ps) or currents (e.p.s.cs) were recorded from the majority of guinea-pig coeliac ganglion neurons after 5 days in culture (Fig. 1A, a). Rise time of the e.p.s.p., measured as the time from 10 to 90% of peak amplitude, was 5.4 ± 2 ms ($n=3$) and the time from peak to half-decay was 22 ± 4 ms ($n=6$); e.p.s.c. rise time was 3.2 ± 0.8 ms, with half-decay of 11 ± 5 ms ($n=5$). These values are similar to those reported for fast synaptic potentials mediated by glutamic acid or acetylcholine (ACh) in other neurons. Suramin is a selective antagonist at P₂-purinoceptors in vascular and visceral smooth muscles^{7–11} and does not block nicotinic e.p.s.ps¹². Suramin reversibly reduced the e.p.s.p. or e.p.s.c. (Fig. 1A and C); concentration required to produce 50% inhibition (IC₅₀) of the synaptic potential was $1.5 \mu\text{M}$ (Fig. 1C). The e.p.s.p. or e.p.s.c. was also inhibited or abolished after desensitization of the P_{2X}-purinoceptor by α,β -methylene-ATP^{1,7,13} ($10 \mu\text{M}$; Fig. 2a, $n=4$).

The e.p.s.p. and e.p.s.c. were the result of synaptically released transmitter because they were reversibly abolished by cadmium ($100\text{--}200 \mu\text{M}$) and/or a low-calcium, high-magnesium external solution (Fig. 2c, $n=5$) and also by tetrodotoxin (TTX, 300 nM ; Fig. 2d, $n=3$). Synaptic responses were unaffected by the nicotinic receptor antagonists hexamethonium ($100 \mu\text{M}$) and curare ($20\text{--}50 \mu\text{M}$), or the 5-hydroxytryptamine (5-HT₃) receptor antagonist ICS 205-930 ($0.5 \mu\text{M}$), when applied alone or in combination (Fig. 2b; $n=4\text{--}6$). The β -adrenoceptor blocker, propranolol ($1 \mu\text{M}$) and the α -receptor antagonist,

FIG. 1 Inhibition of e.p.s.p. (A, a), e.p.s.c. (A, b) and ATP-induced current (B) by suramin. A, a, Superimposed tracings of e.p.s.p. recorded using an intracellular microelectrode from one neuron. Traces show control, 0.3, 1, 3 and $10 \mu\text{M}$ suramin, and 6 min after suramin washout; each is the average of 10 e.p.s.ps evoked at 0.1 Hz. Holding potential -120 mV . A, b, Superimposed e.p.s.cs recorded using whole-cell patch pipette at holding potential of -110 mV . Traces show control, $30 \mu\text{M}$ suramin and 5 min after suramin washout; each is the average of four e.p.s.cs. Filled triangle indicates nerve stimulation (ns). The 50-ms calibration applies to A, a, B. Superimposed currents recorded with whole-cell pipette in response to a roughly 2-s fast flow application of $10 \mu\text{M}$ ATP; traces are control, 1, 10, $30 \mu\text{M}$ suramin and 3 min after wash of suramin; holding potential -70 mV . C, Inhibition of synaptic response (●) and ATP-current (○) in coeliac neurons. ATP current is that evoked by 10 or $30 \mu\text{M}$ ATP at -70 mV ; points are mean $\pm$ s.e.m., with $n=3\text{--}6$ for each point.

METHODS. Methods of dissociation of coeliac ganglia from young adult guinea-pigs and culture conditions have been described²¹. Over 95% of cells were positive for immunohistochemical localization of tyrosine hydroxylase, a marker for catecholamine-containing sympathetic neurons²² ($n=230$ cells, counted from five preparations at day 7 in culture). Intracellular and whole-cell recordings were obtained from coeliac ganglion neurons maintained in culture for 4 to 12 days. Intracellular recordings were made with 1 M KCl-filled microelectrodes ($15\text{--}30 \text{ M}\Omega$) using standard techniques. Physiological saline of (mM) NaCl 126, NaH₂PO₄ 1.2, MgCl₂ 1.2, CaCl₂ 2.5, KCl 5, NaHCO₃ 25, glucose 11, pH 7.25; gassed with 95% O₂ and 5% CO₂ flowed at 3 ml min^{-1} through a small-organ bath (volume 0.5 ml); all drugs were applied by superfusion and temperature was maintained at $33\text{--}34^\circ\text{C}$. A glass electrode with tip diameter of roughly $40 \mu\text{m}$, filled with physiological saline, was used to apply focal extracellular stimulation to fibre networks; stimulus durations ranged from 0.01–0.5 ms and voltage settings on a Grass S88 stimulator ranged from 4–80 V. Standard gigaseal whole-cell recording techniques were used²³; patch pipettes had resistances of 1.5–3 M Ω . Internal solution (mM) was KCl or potassium gluconate (150), EGTA (10), HEPES (10), Mg-ATP (2), GTP (0.1), pH 7.25, except for experiments in which reversal potentials of the e.p.s.c. were measured, in which case caesium gluconate (150) was used. All membrane potentials are corrected for the liquid junction potential. In experiments in which ATP and GTP were omitted from the internal solution, 2 mM MgCl₂ was added. Atropine ($1 \mu\text{M}$) was added to external solution to block muscarinic receptors in all experiments in which ACh was applied. The external solution and focal external stimulation was the same as for intracellular recordings; most experiments were carried out at $23\text{--}25^\circ\text{C}$. Drugs were applied from a multibarrelled 'fast flow' delivery system²⁴; delay to onset of action was 200–800 ms and effects reached equilibrium concentrations within 1–1.5 s. All antagonists and α,β -methylene-ATP were present for 2–4 min before agonist applications.

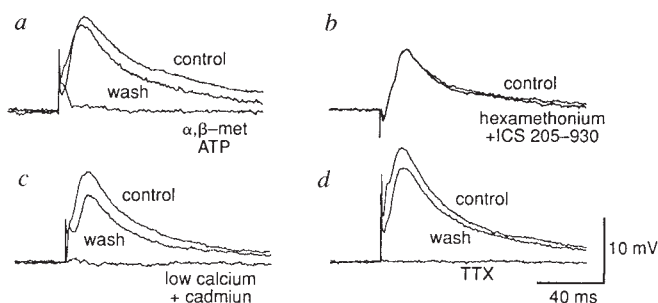


FIG. 2 Synaptic potentials recorded from a coeliac ganglion neuron using an intracellular microelectrode. Each set of superimposed traces show control, drug application and washout (only control and drug in *b*). *a*, α,β -Methylene-ATP (α,β -metATP; 10 μM) depolarized the neuron by 25 mV and this depolarization desensitized by 60% within 2 min, at which time the e.p.s.p. was abolished; e.p.s.p. was recorded at the same membrane potential as control and wash. *b*, The e.p.s.p. was not altered by hexamethonium (100 μM) and ICS 205-930 (0.5 μM). *c*, *d*, Low calcium (0.25 mM), high magnesium (6 mM) solution containing 100 μM cadmium (*c*) or TTX (*d*) also abolished the synaptic potential. Each trace is the average of 10 responses at 0.1 Hz; stimulation parameters were 0.1 ms duration per 40 V for responses in *b*, and 0.05 ms per 40 V for responses in *a*, *c* and *d*. Membrane potential was held at -120 mV to prevent antidromic action potentials; however, in *a* and *c*, an antidromic prepontential precedes the synaptic potential.

phenolamine (1 μM) did not inhibit the e.p.s.p. or e.p.s.c. ($n=4$); however, phenolamine and the selective α_2 -adrenoceptor antagonist, idazoxan (100 nM), significantly increased the amplitude of the synaptic response (23–48% increase in three cells). This result is probably due to blockade of presynaptic α_2 -receptors which are present on sympathetic nerve terminals^{14,15}. Activation of these receptors by synaptically released noradrenaline inhibits the release of many transmitters^{14,15}; this result implies that α_2 -adrenoceptors also inhibit ATP release.

The e.p.s.p. or e.p.s.c. was also unaffected by bicuculline (10 μM ; $n=2$), 6-cyano-2,3-dihydroxy-7-nitro-quinoline (CNQX, 10 μM ; $n=2$), 2-amino-5-phosphonopentanoic acid (APV, 50 μM) and 8-cyclopentyltheophylline (1 μM ; $n=2$), which would be expected to block GABA_A, non-NMDA, NMDA and adenosine receptors. ATP or α,β -methylene-ATP evoked an inward current at the normal resting potential (about -60 mV); at -70 mV, the peak current evoked by ATP (10 μM) was 490 ± 69 pA ($n=5$) and by α,β -methylene-ATP (10 μM) was 510 ± 95 pA ($n=4$). Currents evoked by both agonists were reversibly inhibited by suramin (Figs 1*B*, *C* and 4*a*); IC₅₀ value for inhibition of the ATP current was 2.7 μM (Fig. 1*C*).

The current-voltage relationship for the e.p.s.c. was the same as that for the ATP-induced current (Fig. 3). Both currents showed marked inward rectification at negative potentials (Fig. 3*c*), similar to previously described ATP-induced currents in other neurons^{4–6}. The reversal potential for both the e.p.s.c. and the ATP current was about 5 mV (Fig. 3*c*). Reducing external sodium concentration from 150 to 25 mM shifted the reversal potential for the ATP current to -41 ± 6 mV ($n=3$). Similar results were obtained regardless of whether the internal solution was 150 mM potassium gluconate ($n=4$) or 150 mM caesium gluconate (Fig. 3). The reversal potential for the ATP current

was not shifted when external chloride was replaced with gluconate ($n=2$). Both e.p.s.c.s ($n=2$) and ATP-induced currents ($n=5$) were recorded for up to 2 h when GTP and ATP were omitted from the internal pipette solution. These results show that the e.p.s.c. and the ATP current are due to a ligand-gated cation conductance increase.

Virtually all mammalian autonomic neurons possess both nicotinic and 5-HT₃ receptors which activate a fast cation-selective conductance similar to that produced by ATP. In this study, over 95% of coeliac neurons showed inward currents to ATP ($n>60$), 60% responded to 5-HT and 45% to acetylcholine. To test further the specificity of action of suramin and α,β -methylene-ATP, we compared ATP responses with those evoked by 5-HT and ACh in those neurons in which all agonists produced a response ($n=6$; Fig. 4). Each of these agonists caused an inward current at -70 mV, but only the ATP current was blocked by suramin (Fig. 4). ACh and 5-HT activated nicotinic and 5-HT₃ receptors because they were blocked by hexamethonium and ICS 205-930, respectively ($n=5$; Fig. 4). As for the evoked e.p.s.c., neither hexamethonium nor ICS 205-930 altered the ATP-induced current ($n=6$). Additionally, during superfusion with α,β -methylene-ATP, the ATP-induced current but not the ACh or 5-HT current was abolished ($n=3$). In three neurons in which nerve stimulation evoked an e.p.s.c., ATP (3 μM), but not ACh or 5-HT (3 μM), evoked an inward current.

Taken together, the results strongly indicate that the synaptic and the ATP-induced currents recorded in the present study are due to activation of the same P₂-purinergic receptor. ATP acts as a neurotransmitter at several junctions between autonomic nerves and visceral or vascular smooth muscle^{1,16,17}; indeed, in the guinea-pig the coeliac neurons that innervate the gastrointestinal blood vessels release ATP onto P_{2X}-purinoceptors on vascular smooth muscle to constrict the blood vessel¹². This study

FIG. 3 Amplitude of e.p.s.c. (*a*) and ATP-evoked current (*b*) recorded from one neuron with whole-cell pipette at membrane potentials between -110 mV and 70 mV as indicated; recording made with caesium gluconate-filled pipette. Filled triangle indicates nerve stimulation (ns); bar above upper trace in *b* indicates ATP (3 μM) application. Each e.p.s.c. is the average of six responses to stimulation (0.1 Hz, 0.5 ms per 20 V). *c*, Current-voltage relation for e.p.s.c. (●) and ATP current (○); responses are normalized so that current evoked in individual cells at -110 mV = -1 . Each point is mean $\pm$ s.e.m., $n=4$ for e.p.s.c. and 7 for ATP current.

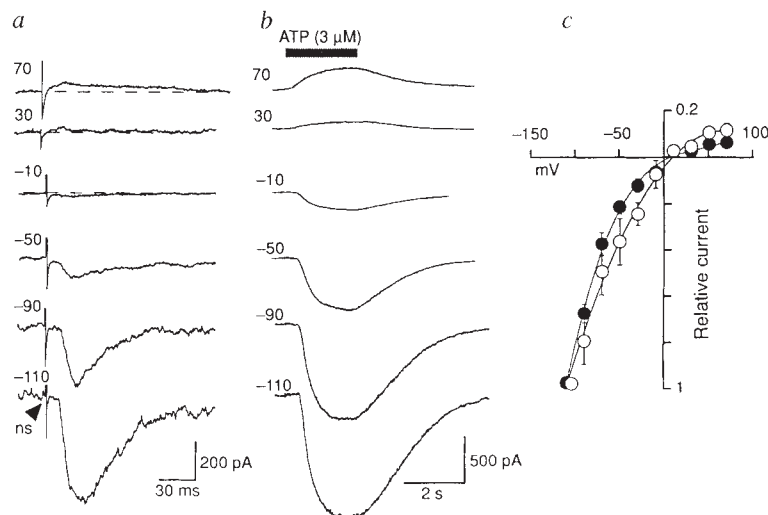
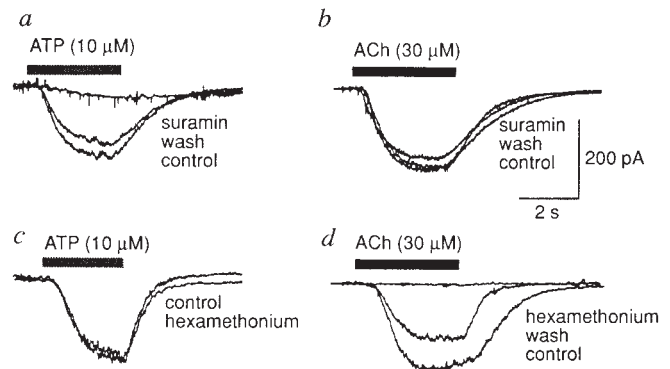


FIG. 4 Suramin selectively inhibits the ATP-induced current (a,b) but hexamethonium selectively blocks the ACh-induced current (c,d). Superimposed traces of whole-cell recordings obtained from one neuron in response to roughly 2-s application of ATP (10 μ M, a, c) or ACh (30 μ M, b, d). Concentrations of ATP and ACh were chosen to give roughly equivalent currents at -70 mV. Similar results were obtained with 5-HT: suramin (30 μ M) did not alter the current evoked by 30 μ M 5-HT ($n=3$). The 5-HT₃ receptor antagonist, ICS 205-930 (100 nM) blocked the 5-HT potential but not the ATP-induced current ($n=3$; data not shown).



has demonstrated that ATP can act as a fast excitatory synaptic transmitter at nerve-nerve synapses. The wide distribution of ATP in synaptic vesicles in the central and peripheral nervous system^{13,18-20} and the presence of ligand-gated P₂-purinoceptors on peripheral and central neurons indicates that this fast transmitter role for ATP may be more widespread than previously realized. □

Received 16 March; accepted 28 April 1992.

1. Burnstock, G. *Nucleosides and Nucleotides* **10**, 917-930 (1991).
2. Jahr, C. E. & Jessel, T. M. *Nature* **304**, 730-733 (1983).
3. Benham, C. D. *et al.* *J. Physiol.* **387**, 473-488 (1987).
4. Bean, B. P. *J. Neurosci.* **10**, 1-10 (1990).
5. Bean, B. P., Williams, C. A. & Ceelen, P. W. *J. Neurosci.* **10**, 11-19 (1991).
6. Fieber, L. A. & Adams, D. J. *J. Physiol.* **434**, 239-256 (1991).
7. Dunn, P. M. & Blakeley, A. G. H. *Br. J. Pharmac.* **93**, 243-245 (1988).

8. Nakazawa, K. *et al.* *Br. J. Pharmac.* **101**, 224-226 (1990).
9. von Kugelgen, I., Bultmann, R. & Starke, K. *Naunyn-Schmiedeberg's Archs Pharmac.* **340**, 760-763 (1989).
10. Hoyle, C. H. V., Knight, G. E. & Burnstock, G. *Br. J. Pharmac.* **99**, 617-621 (1990).
11. von Kugelgen, I., Bultmann, R. & Starke, K. *Naunyn-Schmiedeberg's Archs Pharmac.* **342**, 198-205 (1990).
12. Evans, R. J. & Surprenant, A. *Br. J. Pharmac.* (in the press).
13. Burnstock, G. *Neurochem. Int.* **2**, 357-368 (1990).
14. Starke, K. *Rev. Physiol. Biochem. Pharmac.* **107**, 73-146 (1987).
15. Starke, K., Gothert, M. & Kilbinger, H. *Physiol. Rev.* **69**, 864-969 (1989).
16. Burnstock, G. *Acta Physiol. Scand.* **126**, 67-91 (1986).
17. White, T. D. *Pharmac. Ther.* **38**, 129-168 (1988).
18. Volkand, W. & Zimmermann, H. *J. Neurochem.* **47**, 1449-1462 (1986).
19. Lagercrantz, H. *Neuroscience* **1**, 81-92 (1976).
20. Fried, G. *Acta. physiol. scand.* **493**, 546-559 (1980).
21. Coggan, J. S. & Kreulen, D. L. *J. Auton. Nerv. Syst.* **34**, 147-156 (1991).
22. Alm, P. *et al.* *Neuroscience* **4**, 145-158 (1979).
23. Hamill, O. P. *et al.* *Pflügers Arch.* **391**, 85-100 (1981).
24. Shen, K.-Z. *et al.* *J. Physiol.* **431**, 585-608 (1990).

ACKNOWLEDGEMENTS. We thank M. Meyers for dissociation and tissue culture procedures and R. A. North for comments. This work supported by grants from the NIH.

SWI6 protein is required for transcription of the periodically expressed DNA synthesis genes in budding yeast

Noel F. Lowndes*, Anthony L. Johnson*, Linda Breeden† & Leland H. Johnston*‡

* Laboratory of Yeast Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

† Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104, USA

IN budding yeast many genes are expressed under cell-cycle control in late G1. These include a large group of DNA synthesis genes¹, the *HO* gene² involved in mating-type switching, *CTS1* (chitinase)³ and also *CLN1* and *CLN2* (ref. 4) encoding G1 cyclins. Two factors, encoded by the *SWI4* and *SWI6* genes, are required for *HO* (ref. 5), *CLN* (refs 6, 7) and *CTS1* (ref. 3) gene expression and, at least in the *HO* promoter, bind to CACGA₄ upstream sequences^{5,7-9} (CCBs). This motif is not found upstream of the DNA synthesis genes, which instead have a hexamer element, ACGCGT¹ (MCB), an *MluI* restriction site, that is recognized by a cell-cycle regulated transcription complex DSC1 (ref. 1). This *MluI*-activation system consisting of the MCBs and DSC1 is conserved in fission yeast where a DSC1-like complex controls the *cdc22⁺* ribonucleotide reductase gene¹⁰. The *Schizosaccharomyces pombe cdc10⁺* gene encodes a component of DSC1 (ref. 10) and, significantly, this has homology with both the Swi4

and Swi6 proteins^{8,11}. Here we show that Swi6 is an essential component of DSC1 and that deletion of *SWI6* impairs the cell-cycle regulation of the DNA synthesis genes, as well as *CLN1* and *CLN2*. Thus Swi6 is the common factor in regulation of all the above genes and may therefore be responsible for the timing of their expression in late G1.

The DSC1 transcription factor was identified as an activity that binds to synthetic copies of the MCB sequences in gel-retardation assays¹. The presence of DSC1 was therefore examined by this means in *swi4* and *swi6* mutants, and whereas *swi4* mutants were without effect, *swi6* mutants abolished detectable DSC1 activity (Fig. 1a, lanes 2-5). Introduction into a *swi6* mutant of the *SWI6* gene on a plasmid restored DSC1 (Fig. 1a, lane 6). By contrast, introduction of the fission yeast *cdc10⁺* gene into a *swi6* deletion mutant did not restore DSC1 even when *cdc10⁺* was expressed from the powerful budding yeast *GAL1* promoter (Fig. 1a, lanes 7-9). Thus *SWI6* either controls or is a component of DSC1 but, in either case, *cdc10⁺* cannot substitute for its function *in vitro*.

When affinity-purified antibodies against Swi6 protein were added to a binding reaction there was a super-shift of DSC1 to form a ternary complex of antibody-bound DSC1 (Fig. 1b, lanes 3-5). By contrast, the addition of non-specific serum had no effect on retardation of DSC1 (data not shown). As the concentration of anti-Swi6 antibodies was increased the level of DSC1 and the ternary complex was progressively reduced in contrast to the other retarded species, emphasizing that the antibodies were interacting specifically with DSC1. In addition, when the Swi6 protein was increased in size by the addition of amino-terminal amino acids from *GAL4*, there was an increased retardation of DSC1 (data not shown). The Swi6 protein is thus a component of DSC1 rather than an activity that controls it indirectly.

As Swi6 is a component of DSC1, deletion of *SWI6* might be expected to affect expression of a gene driven by ACGCGT

‡ To whom correspondence should be addressed.

sequences. Synthetic copies of this sequence in an appropriate test plasmid stimulate expression of β -galactosidase¹ and this expression is indeed abolished by *swi6* (Table 1, line 4). By contrast, a *swi4* deletion had no significant effect on β -galactosidase levels from this plasmid (Table 1, line 4). An additional 3' purine two nucleotides downstream of the T residue has been implicated in control of the DNA synthesis genes^{12,13} and expression driven by synthetic copies of this extended sequence are also affected by *swi6* (Table 1, line 5). Even expression stimulated by a 477-base-pair (bp) promoter fragment from *CDC9* (DNA ligase, one of the cell-cycle-regulated DNA synthesis genes) is substantially diminished in a *swi6* mutant, although in this case there was some 20% residual expression (Table 1, line 6). These data are consistent with the Swi6 protein being part of a complex that binds to and stimulates transcription from the MCB elements. Surprisingly, in view of this, deletion of *SWI6* has little effect on mid-log phase transcript levels from the DNA synthesis genes (see below). One explanation for this discrepancy is that without the Swi6 protein, the remaining components of DSC1 may still bind to MCBs (although this complex would have to be unstable as there is no evidence for it in our band shift assays; see Fig. 1a, lanes 4 and 5) and enhance transcription from the DNA synthesis genes. But when bound to the MCBs inserted into pLGΔ178, this residual DSC1 may be unable to interact productively with the cytochrome *c* minimal promoter in pLGΔ178. Nonetheless, the effect of the *swi6* mutation on β -galactosidase expression from pLGΔ178.3M does emphasize its role in the *MluI*-activation system.

RNA hybridization analysis showed that deletion of the *SWI6* gene had no significant effect on mid-log phase levels of transcripts from the DNA synthesis genes (data not shown) but, importantly, this deletion did affect their cell-cycle regulation.

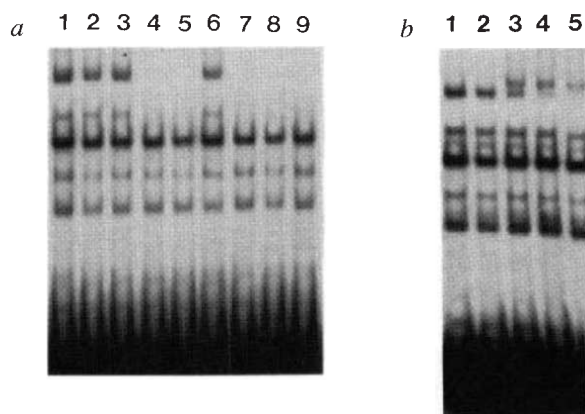


FIG. 1 Gel retardation analysis showing that the Swi6 protein is a component of DSC1. *a*, Extracts from *swi6* mutants are defective in DSC1. DSC1 activity corresponds to the upper band which is the specific retarded species¹. Assays were carried out with clarified crude extracts prepared from the following strains: lane 1, BY28, a derivative of W303-1A (ref. 6) that is congenic with the *swi4* and *swi6* mutants; lane 2, BY136, containing *swi4-100*, a point mutation⁵; lane 3, BY604, a *swi4* deletion¹⁴; lane 4, BY142, containing *swi6-399*, a point mutation⁵; lane 5, BY600, a *swi6* deletion¹¹; lane 6, BY600 containing a plasmid with the wild-type *SWI6* gene; lane 7, BY600 containing plasmid pDB262-*cdc10*⁺ carrying the *S. pombe cdc10*⁺ gene; lane 8, BY600 with plasmid pGAL-*cdc10*⁺ expressing *S. pombe cdc10*⁺ under control of the *GAL1* promoter and grown in medium containing galactose; lane 9, the same strain grown in medium containing glucose. *b*, DSC1 contains the *SWI6* gene product. Binding reactions were allowed to equilibrate as above and then further additions were as follows: lane 1, no additions; lane 2, binding buffer; lanes 3–5, 2 μ l, 4 μ l and 6 μ l, respectively, of affinity-purified anti-Swi6 serum.

METHODS. All binding reactions were done with 0.5 ng of a [γ -³²P] 5' end-labelled 73-bp *AluI*-*Sau3A* fragment containing three synthetic ACGCGT sequences at $\sim 1 \times 10^8$ c.p.m. μ g⁻¹ and 10 μ g crude protein extract was used¹. The preparation of clarified crude extracts, protein estimations, binding buffer and gel running conditions were as described previously¹.

TABLE 1 The effect of *swi6* mutation on β -galactosidase expression under control of ACGCGT sequences

Construct	Activating sequence	β -Galactosidase level		
		SWI+	<i>swi6</i> ::TRP1	<i>swi4</i> ::LEU2
pLGΔ312	UASc	15.05	16.24	17.36
pLGΔ178	—	0.05	0.06	—
pLGΔ178.3mut	3 × ActaGT	0.06	0.06	—
pLGΔ178.3M	3 × ACGCGT	4.00	0.08	3.68
pLGΔ178.3MX	3 × ACGCGTTAA	5.05	0.16	—
pLGΔ178.477	<i>CDC9</i> promoter fragment	1.50	0.33	—

pLGΔ312 and pLGΔ178 are vectors carrying elements of the budding yeast cytochrome *c* promoter fused to the *E. coli lacZ* gene²⁰. The former includes the entire *CYC1* promoter, including the upstream activating sequence (UASc), and the latter contains only the TATA region without UASc. The constructs pLGΔ178.3 mut and pLGΔ178M have been described previously¹. pLGΔ178.3MX was constructed by inserting three repeats of the double-stranded oligonucleotide 5'-TCGAGACGCGTTAAC-3'/5'-TCGAGTTAAGCGTGC-3' in the direct orientation into the *XhoI* site of pLGΔ178. pLGΔ178.477 was constructed by inserting the 477-bp *AluI* fragment from the *CDC9* promoter from positions -65 to -542 relative to the ATG²¹ into the pLGΔ178 *XhoI* site. All of these constructs were transformed into congenic wild-type, *swi4* and *swi6* strains (BY28, BY604 and BY600, derivatives of W303-1A) and mid-log phase cells were assayed for β -galactosidase. The results are expressed as absorbance at 420 nm per 250 μ g of crude cell protein per hour and each value is the average of assays on three independent transformants.

A culture was synchronized by holding a *swi6 cdc15*^{ts} double mutant at 37 °C and then returning the cells to 25 °C (*swi6* cells are too asymmetric for elutriation and they do not release synchronously from an α -factor induced cell cycle block). This procedure produced reasonable synchrony but it decayed rapidly after one cell cycle. As a control we synchronized the same strain containing a plasmid carrying the wild-type *SWI6* gene. Both cultures gave comparable levels of synchrony as indicated by bud counts and periodicity of histone H2B (Fig. 2a). There was detectable cell-cycle regulation of *CDC9* in the *swi6* delete, but it was markedly less than in the control (Fig. 2b). For instance, the control transcript level increased 2.3-fold after 1 h at 25 °C and this was followed by a threefold decrease, whereas the comparable figures for the *swi6* cells were 1.4-fold for both. A similar result was obtained with *POL1*, DNA polymerase I (data not shown). With *RNR1*, encoding the regulatory subunit of ribonucleotide reductase, the periodic expression was dramatically reduced in *swi6* cells (Fig. 2b). The transcript in the control culture increased no less than 20-fold after 1 h at 25 °C without there being any increase at this time in the *swi6* culture, although a small peak was apparent at 2 h 15 min (Fig. 2b). Deletion of *swi6* therefore severely impairs *CDC9* and *POL1* periodic expression and completely eliminates the correctly timed expression of *RNR1*.

SWI6 is also involved in the control of *CLN1* and *CLN2* (refs 6, 7) and its deletion also affected their cell-cycle regulation (Fig. 2c). As for *CDC9* and *POL1*, periodicity of the *CLNs* was not eliminated but substantially reduced. The increase in the *CLN1* transcript in *swi6* cells after 1 h at 25 °C was no more than 1.8-fold, followed by a similar decrease, whereas the comparable values for the *SWI*⁺ cells were 15-fold and threefold, respectively. With *CLN2*, the *swi6* deletion simply reduced the transcript level without affecting the periodicity a great deal. It is surprising that *SWI6* deletion eliminates *HO* expression⁵ but has so little effect on *CLN1* and *CLN2*, where it is also believed to act on CACGA₄ sequences^{6,7}. Conceivably, the *CLNs* could be under dual regulation, *CLN1* contains a *MluI* sequence in its upstream together with two near-matches, whereas *CLN2* has three *MluI* near-matches⁷.

Finally, we examined the level of the *HO* gene transcript in

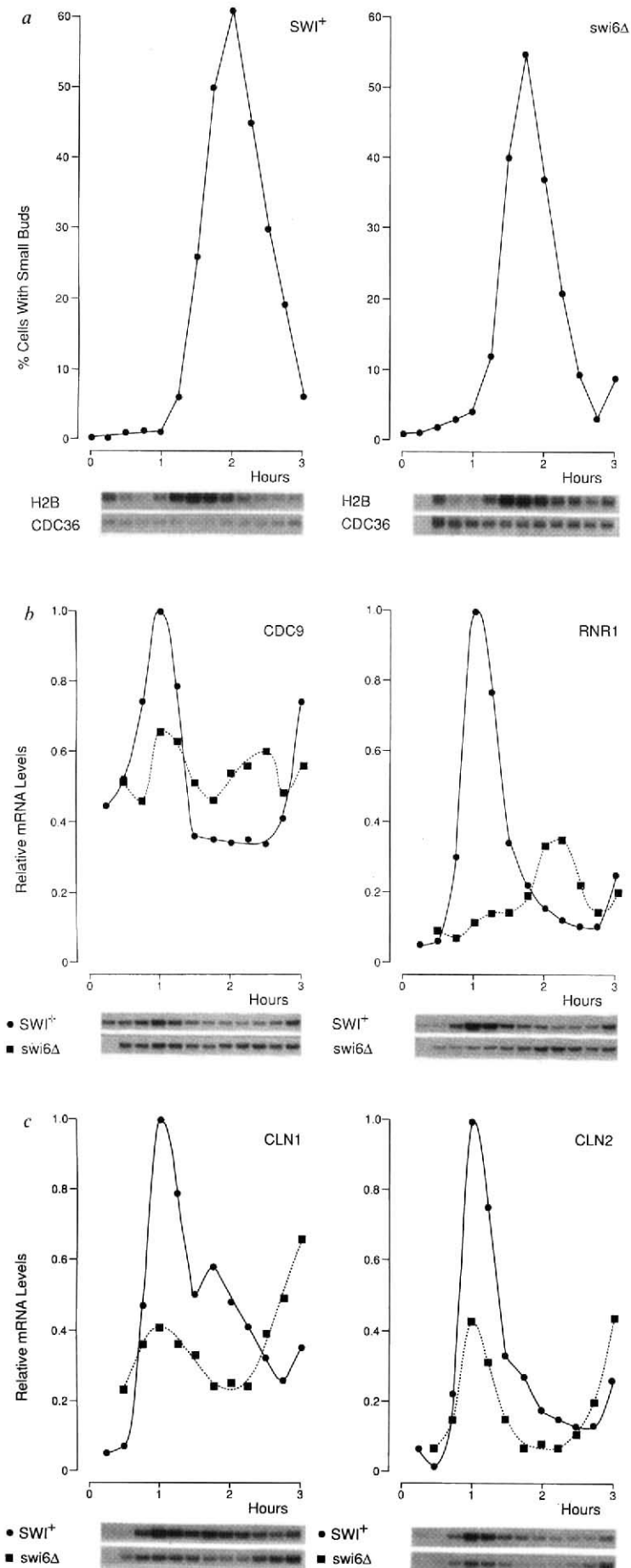
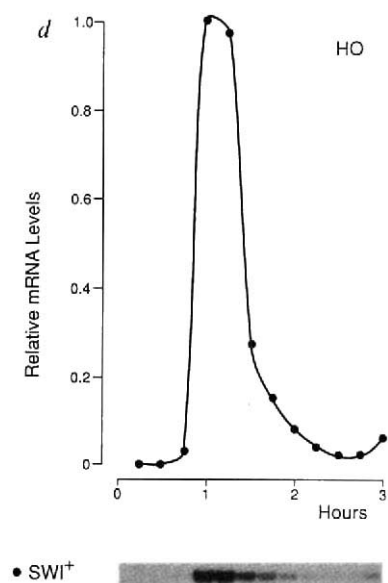


FIG. 2 The cell-cycle regulation in a *swi6* deletion mutant of genes expressed in G1. **a**, Cell-cycle parameters in the *SWI⁺cdc15^{ts}* and *swi6 cdc15^{ts}* synchronous cultures. The budding profile and periodicity of the histone H2B indicate the degree of synchrony obtained and the *CDC36* transcript is a loading control for the RNA blot as it is invariant in the cell cycle. Transcript levels in the above cultures were also determined for: **b**, the DNA synthesis genes *CDC9* and *RNR1*; **c**, the cyclins *CLN1* and *CLN2*; **d**, the *HO* gene. RNA hybridization (northern blot) analysis was carried out on cells removed at the intervals shown, for cultures synchronized by holding the strains at 37 °C followed by transfer to 25 °C. The curves in **b**, **c** and **d**, were obtained by densitometry of the autoradiographic data shown below each figure. The RNA derived from both the *SWI⁺* and *swi6* cultures was electrophoresed on one gel and transferred to a single membrane so that RNA levels are directly comparable. Note that in the *swi6* culture there is no 15-min sample. METHODS. The strain L211-11C *MATa cdc15 swi6::TRP1 trp1 ura3 leu2 his⁻ ade⁻* was constructed by crossing BY600 (ref. 11) with an appropriate strain carrying the *cdc15-1* mutation. This strain, L211-11C, with or without the plasmid containing the wild-type *SWI6* gene, was grown to mid-log phase at 25 °C in yeast nitrogen-base medium²². After sampling, cultures were transferred to 37 °C for 3 h before return to 25 °C. Sampling was continued every 15 min thereafter. RNA extraction, northern blotting and probing have been described²².



the *SWI*⁺ and *swi6* synchronous cultures (Fig. 2d). As expected for an authentic *swi6* mutation⁵, *HO* expression was abolished in the *swi6* culture. But the *HO* transcript was sharply periodic in the *SWI*⁺ culture and, at least in this experiment, the peak occurred at the same time as the two DNA synthesis genes and *CLN1* and *CLN2*.

The perturbation in expression of the *CLNs* and the DNA synthesis genes in a *swi6* mutant might be expected to affect the cell cycle and, in fact, *swi6* strains do grow more slowly than an isogenic wild type. Flow cytometry suggested that the slow growth might be due to a protracted S phase (data not shown) which could reflect an imbalance in the replication proteins. But given the total number of genes regulated by *SWI6*, its deletion might be expected to be far more deleterious, although there is of course residual cell-cycle regulation in at least some of the genes. The residual regulation might indicate the presence of another activity able to substitute for Swi6 to some degree (for a more detailed discussion, see ref. 23). Alternatively, in the case of *CLN1* and *CLN2*, the Swi4 protein may be responsible for the residual cell-cycle regulation. *SWI4* is expressed under cell-cycle control and this expression is essential for the cell-cycle regulation of *HO*¹⁴. Thus an analogous protein to Swi4 in DSC1 might account for the residual expression of the DNA synthesis genes.

There certainly must be at least one other component in DSC1 apart from Swi6. When purified from *Escherichia coli*, Swi6 did not bind specifically to MCB elements (data not shown) and, in addition, the sequence of these elements differs from that of the CCB elements. Both facts point to another protein being involved in determining sequence specificity. The occurrence of different systems for expressing genes in late G1 must have some physiological basis, possibly in the precise timing of gene expression. The role of Swi6 in the regulation of these diverse genes may be to sense cell-cycle position because, so far, it is the only factor that is common to the regulation of *HO*⁵, *CTSI*³, *CLN1* and 2 (refs 6, 7) and the DNA synthesis genes (this letter). The mechanism by which Swi6 might sense position is not known, but one possibility is phosphorylation by p34^{cdc2} (encoded by *CDC28* in budding yeast). Swi6 is a phosphoprotein (L.B., unpublished observation), it has potential p34^{cdc2} kinase phosphorylation sites and it can be phosphorylated *in vitro* by p34^{cdc2} (N.F.L., C. Norbury and L.B., unpublished observation). Furthermore, it is homologous to the fission yeast *cdc10*⁺ gene product¹¹ which is a phosphoprotein¹⁵ and is a component of a DSC1-like transcription factor in *S. pombe*¹⁰. By implication, Cdc10 may be the protein that senses passage through late G1 in fission yeast. But the exact functional relationship between Cdc10 and Swi6 is not clear. Antibodies against the Cdc10 protein had no effect on DSC1 (N.F.L., unpublished observation) and overexpression of *cdc10*⁺ in *Saccharomyces cerevisiae* did not restore DSC1 in *swi6* (Fig. 1a), nor does it suppress the need for Swi6 in *HO* transcription¹¹ (L.B., unpublished observation).

The *SWI6* gene product can be regarded as a general transcription factor for the control of genes expressed in late G1. There may therefore be other cell-cycle stage-specific transcription factors that coordinate the expression of particular groups of genes. For instance *SWI5* (ref. 16) and several *CLB* (refs 17, 18) genes are expressed in G2, and *DBF2*, a putative cell cycle protein kinase, and two of its suppressor genes are coordinately expressed in late M phase or early G1 (ref. 19, and L.H.J., unpublished observation). □

8. Andrews, B. J. & Herskowitz, I. *Nature* **342**, 830–833 (1989).
9. Taba, M. R. M., Muroff, I., Lydall, D., Tebb, G. & Nasmyth, K. A. *Genes Dev.* **5**, 2000–2013 (1991).
10. Lowndes, N. F., McInerney, C. J., Johnson, A. L., Fantes, P. A. & Johnston, L. H. *Nature* **355**, 449–453 (1992).
11. Breeden, L. & Nasmyth, K. A. *Nature* **329**, 651–654 (1987).
12. McIntosh, E., Atkinson, T., Storms, R. K. & Smith, M. *Molec. cell. Biol.* **11**, 329–337 (1991).
13. Johnston, L. H., Lowndes, N. F., Johnson, A. L. & Sugino, A. *Cold Spring Harb. Symp. quant. Biol.* (in the press).
14. Breeden, L. & Mikesell, G. E. *Genes Dev.* **5**, 1183–1190 (1991).
15. Simanis, V. & Nurse, P. *J. Cell Sci.* **92**, 51–56 (1989).
16. Nasmyth, K. A., Seddon, A. & Ammerer, G. *Cell* **49**, 549–558 (1987).
17. Surana, U. et al. *Cell* **65**, 145–161 (1991).
18. Ghilara, J. B. et al. *Cell* **65**, 163–174 (1991).
19. Johnston, L. H., Eberly, S. L., Chapman, J. W., Araki, H. & Sugino, A. *Molec. cell. Biol.* **10**, 1358–1366 (1990).
20. Guarente, L. & Mason, T. *Cell* **32**, 1279–1286 (1983).
21. Barker, D. G., White, J. H. M. & Johnston, L. H. *Nucleic Acids Res.* **13**, 8323–8336 (1985).
22. White, J. H. M., Barker, D. G., Nurse, P. & Johnston, L. H. *EMBO J.* **5**, 1705–1709 (1986).
23. Lowndes, N. F. & Johnston, L. H. *Trends Genet.* **8**, 79–81 (1992).

ACKNOWLEDGEMENTS. We thank J. Sidorova for characterizing the Swi6-specific reagents used in Fig. 1. N.F.L. was partly supported by a grant from the EC and L.B. was funded by grants from the NIH and ACS.

A central role for *SWI6* in modulating cell cycle Start-specific transcription in yeast

Léon Dirick, Thomas Moll, Herbert Auer & Kim Nasmyth*

Institute of Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria

MOST genes involved in DNA replication in the yeast *Saccharomyces cerevisiae* are transcribed transiently during late G1 as cells become committed to a new cell cycle at Start¹. Their promoters all contain one or more versions of an 8-base-pair motif (ACGCGTNA) containing an *MluI* restriction enzyme site and called the *MluI* cell-cycle box (MCB)². MCBs are both necessary and sufficient for the late G1-specific transcription of the *TMP1* thymidylate synthase and *POL1* DNA polymerase genes^{3,4}. A different late G1-specific 8-base-pair transcription element called the SCB (CACGAAAA; ref. 5) is bound by a factor containing the Swi4 and Swi6 proteins^{6,7}. We describe here the formation *in vitro* of complexes on *TMP1* MCBs that contain the Swi6 protein and, we suggest, a protein of relative molecular mass 120,000 (p120) that is distinct from Swi4. Transcription due to SCBs and MCBs occurs in the absence of Swi6 but it is no longer correctly regulated in the cell cycle. We suggest that Swi6 is an essential regulatory subunit of two different Start-dependent transcription factors. One factor (SBF) contains Swi4 and binds to SCBs, whereas the other (MBF) contains the protein p120 and binds MCBs.

Mixing crude extracts with a natural oligonucleotide from the *TMP1* promoter containing two MCBs (MCB-*TMP1*) produces a slow-migrating gel retardation complex that is disrupted by competition with unlabelled oligonucleotide or oligonucleotides containing multiple SCBs but not by an unrelated oligonucleotide (Fig. 1a). The complex is formed by *swi4* but not by *swi6* mutant extracts (Fig. 1a). Treatment with Swi6-specific antibodies abolished complex formation, whereas preimmune and Swi4-specific antibodies had no effect (Fig. 1a). We conclude that the Swi6 protein is part of the complex but that Swi4 is not. We call the binding activity MBF (*MluI* binding factor). There are no major changes in the amount of MBF in a synchronous culture⁷ produced by release from α factor-induced G1 arrest (Fig. 1b). Moreover, the level of MBF in cycling cells and in a population of unbudded G1 cells purified by elutriation is similar (data not shown).

MBF purified 150-fold by heparin-agarose chromatography⁸ protects specifically both MCB sites in the *TMP1* promoter from

* To whom correspondence should be addressed.

Received 20 February; accepted 2 April 1992.

1. Lowndes, N. F., Johnson, A. L. & Johnston, L. H. *Nature* **350**, 247–250 (1991).
2. Nasmyth, K. A. *Nature* **302**, 670–676 (1983).
3. Dohrmann, P. R. et al. *Genes Dev.* (in the press).
4. Wittenberg, C., Sugimoto, K. & Reed, S. I. *Cell* **62**, 225–237 (1990).
5. Breeden, L. & Nasmyth, K. A. *Cell* **48**, 389–397 (1987).
6. Nasmyth, K. A. & Dirick, L. *Cell* **66**, 995–1013 (1991).
7. Ogas, J., Andrews, B. J. & Herskowitz, I. *Cell* **66**, 1015–1026 (1991).

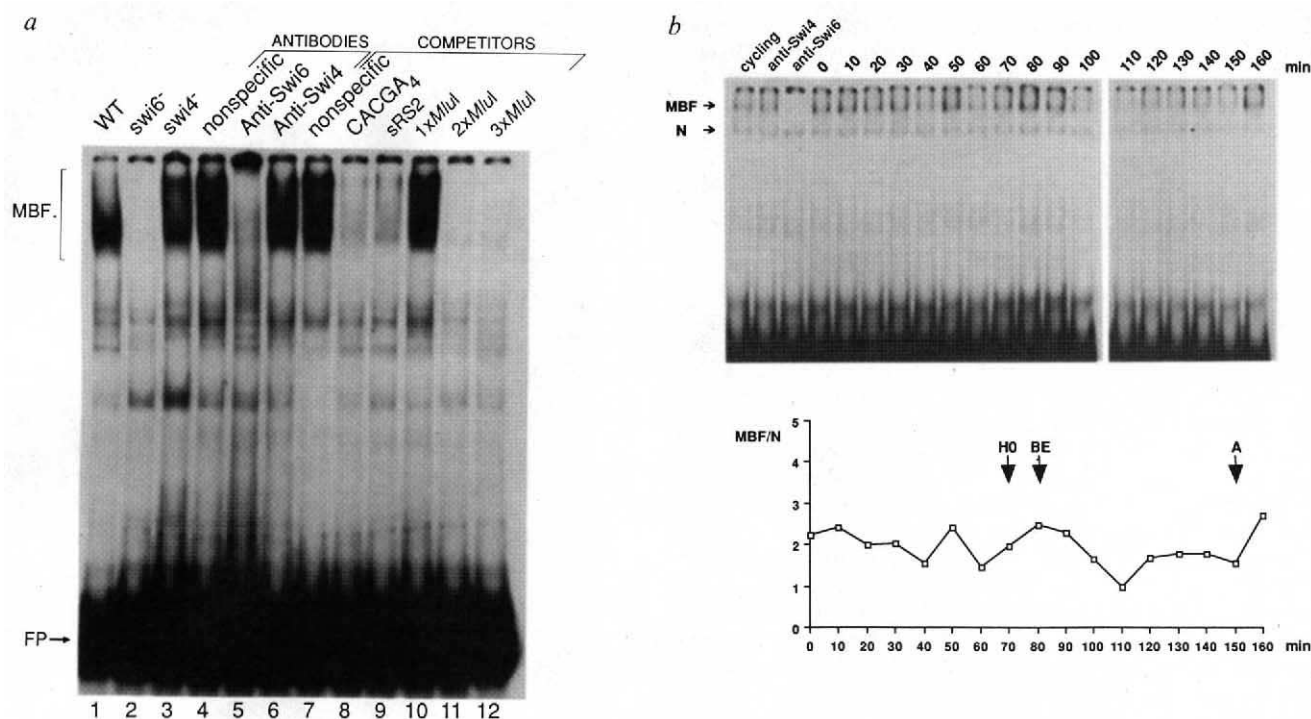


FIG. 1 Swi6 is part of a factor (MBF) that can bind to MCBs at all stages of the cell cycle. **a**, Complexes formed on MCB-TMP1 in crude extracts analysed by gel electrophoresis. Whole-cell extracts were prepared from wild-type (K1268), *swi4* mutant (R/H1071) and *swi6* mutant (K1354) strains⁷. MBF bracket indicates the slow-migrating complex that is present in *swi4* but not in *swi6* mutants (lanes 1–3). Deletion of *SWI4* is lethal for K1268 but such deletions can be obtained if *CLN2* is expressed from the *spADH* promoter. This strain (K2527) also contains wild-type levels of MBF (data not shown). Lanes 4–6 show the effect of adding antibodies (1:20 dilution of sera) directed against dihydrofolate reductase (nonspecific), Swi6 and Swi4 proteins to extracts from the wild type (WT). Lanes 7–12 show the results of competition with non-labelled oligonucleotides. Note that oligonucleotides containing several SCBs as well as MCBs (see Methods) can compete. FP denotes unbound oligonucleotide. **b**, The amount of MBF in α -factor-synchronized cells analysed by gel electrophoresis (top). The graph shows the ratio (by phosphorimaging each lane) of MBF to nonspecific (N) complexes in cycling, α -factor-arrested and released cells (R1793). In this culture, the Start-dependent *HO* gene is fully activated by 70 min, buds emerge (BE) by 80 min, and the frequency of anaphase cells peaks at 150 min (A)⁷. Mixing wild-type and *swi6* mutant extracts showed that the amount of MBF assayed is proportional to its concentration. **c**, MBF binding requires both MCBs. Partially purified MBF was mixed with wild-type (MCB-TMP1) and mutant oligonucleotide probes in the presence or absence of anti-Swi6 antibodies. The 5' and 3' MCBs were mutated from ACGCGTTA to ACTCATTAA and from ACGCGTTT to ACTCATT, respectively (top-strand sequences). Note that both mutations eliminate complexes marked MBF without causing the appearance of faster-migrating Swi6-dependent complexes owing to the binding of MBF to the remaining single site. The mutations appear to create a binding site for a new factor.

METHODS. **a**, The MCB-TMP1 oligonucleotide is a 54-bp fragment from the *TMP1* promoter (–163 to –110 5' of the initiation codon³), with the sequence 5'-TGGTGACGCGTTAAATAGAAAAAATGAAAAAGACCTTAATTGACGCGTTCC-TG-3', flanked by 5' *Sall* and 3' *Bam*HI linkers. Underlined are the two *Mlul* sites. MCB-TMP1 was end-labelled with Klenow polymerase and [α -³²P]dGTP. The oligonucleotides used as competitors are sRS2, from the *HO* promoter⁷, a 53-bp oligonucleotide containing the Swi5 binding site (non-specific)²⁵, an 80-bp sequence containing 5 × TCGATCCACGAAAA (CACGA4), a single *Mlul* site ACGCGTTTAA with *Xho*I linkers (1 × *Mlul*), MCB-TMP1 itself (2 × *Mlul*) and MCB-TMP1 with an additional *Mlul* site, GTGACGCGTTAAAG 3' (3 × *Mlul*). Binding was assayed in 20- μ l reactions as described (ref. 7), except that glycerol was added to 20% (v/v) final concentration. Total protein extracts (20 μ g) were mixed in this buffer with ~0.05 pmol labelled probe, samples were incubated at room temperature for 15 min and loaded onto a 5% polyacrylamide gel (30:1 cross linking) and run for 2 h at 200 V in 0.5 × TBE buffer. Unlabelled competitors were added to a 50-fold molar excess. When antibodies were used, 1 μ l diluted antiserum⁷ (1:20) was

added after 10 min and further incubated for 5 min before loading. Rabbit polyclonal Swi4 antibodies were raised against a polypeptide encoded by an internal *Acl*I restriction enzyme fragment of *SWI4* (sequence given in ref. 6) and will be described elsewhere. They recognize only the Swi4 protein by western analysis of crude extracts and cause a supershift of Swi4/Swi6-SCB gel retardation complexes (data not shown). **b**, The synchronous culture and extracts prepared from it are described in ref. 7. Protein (40 μ g) was mixed with the MCB-TMP1 probe and complexes electrophoresed as in **a**, except that the reaction volume was 30 μ l. MBF complexes, nonspecific complexes (N), and the background in each lane were quantified using phosphorimaging (Molecular Dynamics). The background was subtracted from MBF and N and the ratio of MBF to N was plotted. Anti-Swi4 and anti-Swi6 antibodies were used (lanes 2 and 3) to check the identity of MBF complexes. **c**, Whole-cell extracts from a protease-deficient strain (K1142) were prepared and fractionated over a heparin-agarose column²⁵. Fractions were analysed for the formation of gel retardation complexes with the MCB-TMP1 oligonucleotide. MBF activity eluted from the column at 0.3 M (NH₄)₂SO₄ (data not shown). Partially purified MBF (150-fold) was then used for DNA-binding and gel retardation reactions.

cleavage by copper phenanthroline⁹ (Fig. 2a). We have analysed which G residues interfere with MBF binding when methylated by dimethylsulphate¹⁰ (Fig. 2b) and which A or G residues interfere when carbethoxylated by diethyl pyrocarbonate¹¹ (Fig. 2c). Modification of bases in both MCBs interferes with complex formation, suggesting that MBF binding is cooperative. Consistent with this, oligonucleotides containing only the 5' MCB cannot compete for binding (Fig. 1a) and point mutations in either of the MCBs in MCB-TMP1 abolish complex formation without producing complexes in which only a single MCB is occupied (Fig. 1c). The 26-base-pair separation is not obligatory because MBF also binds to MCBs separated by 16 base pairs (data not shown). The interference pattern at the two MCBs is similar if we assume that both sites are asymmetric and have opposing orientations (Fig. 2d). MBF binding involves two A-T base pairs outside the *Mlu*I restriction site, also necessary for MCB function *in vivo*³. MBF makes contact along an extensive stretch of the major groove of each MCB and must therefore bind to both sides of the double helix at each site.

We have measured the levels of β -galactosidase activity in wild-type (K2811) and *swi6* Δ mutant (K2803) strains containing plasmids expressing *LacZ* either from a *CYC1* promoter lacking upstream activating sequences (pLGA178)¹² or from a derivative containing MCB-TMP1. The *TMP1* upstream activating sequence increases activity from 17 units to 240 units in wild type but only from 14 to 33 units in the mutant. This suggests that MBF-DNA complexes also form *in vivo* and can activate transcription. *TMP1* transcription depends on MCBs in its promoter³, but surprisingly we find no effect of deleting *SWI6* on either the level or the start sites of *TMP1* transcripts (Fig. 3a). We do, however, see effects on the ribonucleotide reductase *RNR1* gene (Fig. 3a), which is transcribed at Start and contains MCBs^{13,14}. Mutation of *SWI6* but not *SWI4* causes a large increase in the number of *RNR1* messenger RNAs that start upstream of the normal cluster of start sites, whose abundance is also reduced.

To test whether *Swi6* is important for regulating rather than activating MCB transcription, we have made congenic *SWI6*⁺ and *swi6* Δ strains whose three G1 cyclin genes *CLN1*, 2 and 3 have been deleted and whose growth is dependent on the conditional expression of *CLN1* from the *GAL1-10* promoter¹⁵. Removal of galactose from the medium causes such cells to arrest in G1 and readdition leads, at least in *SWI6*⁺ cells¹⁵, to their synchronous passage through Start. We used this method because the G1 cyclin deficiency of *swi6* Δ mutants¹⁶ causes slow and asynchronous release from α factor-induced G1 arrest (data not shown). Remarkably, the sharp periodicity of *TMP1* RNAs is abolished by the *swi6* mutation (Fig. 3b). Moreover, the normal (Fig. 3b) and abnormal (data not shown) *RNR1* transcripts are also deregulated. In contrast the periodicity of the G1/S specific histone H2B RNAs and the S/G2 specific *SWI5* (ref. 17) and B-type cyclin *CLB2* (refs 18, 19) RNAs is little affected (Fig. 3b, and data not shown). The cell-cycle periodicity of G1 cyclin transcription dependent on SBF^{16,20} is either reduced (*CLN2*) or abolished (*CLN1*) by the *swi6* mutation (Fig. 3b).

SBF binding to SCBs is due to a second subunit, *Swi4*, which has a site-specific DNA-binding domain and can recruit *Swi6* into ternary complexes (M. Primig, manuscript submitted). To determine whether MBF also has a second subunit that can direct binding to MCBs, we analysed which proteins can be crosslinked to 5-bromodeoxyuridine (BrdU)-substituted DNA after ultraviolet irradiation of gel-purified MBF-MCB complexes²¹. A single protein of *M_r* ~ 120K (p120) is detected, whose recovery is greatly reduced by adding *Swi6*-specific antibodies to the binding reaction (Fig. 4a). The gel shown in Fig. 4a was rehydrated and *Swi6* detected by western blotting. *Swi6* from a crude extract and from the crosslinked DNase I-treated complexes runs alongside the 92.5K marker, well in front of p120. *Swi4* is predicted to be 123K (ref. 6), but it cannot be p120

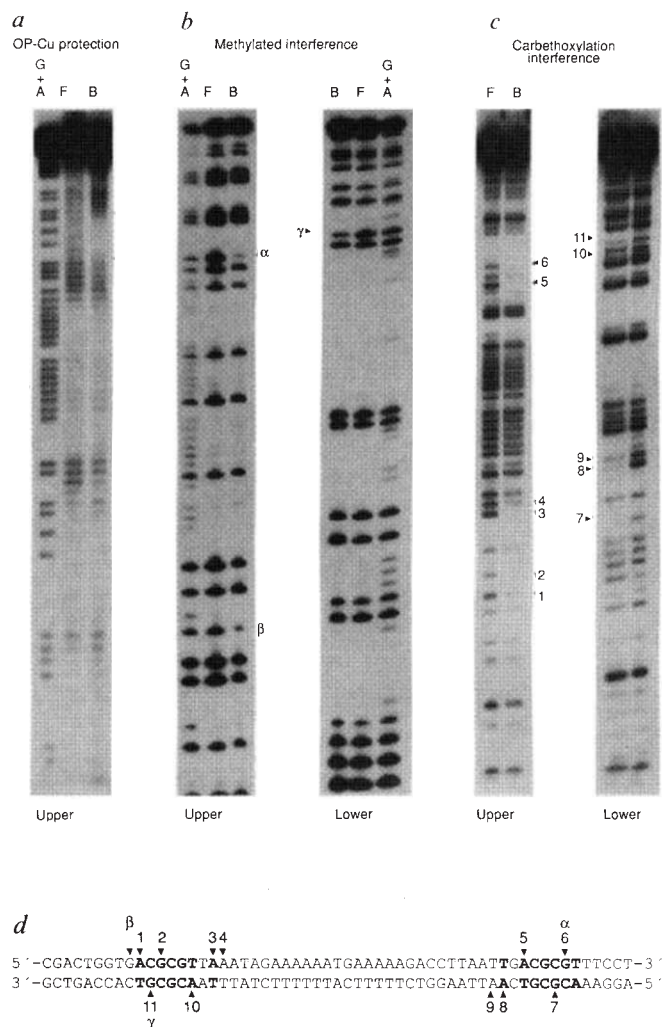
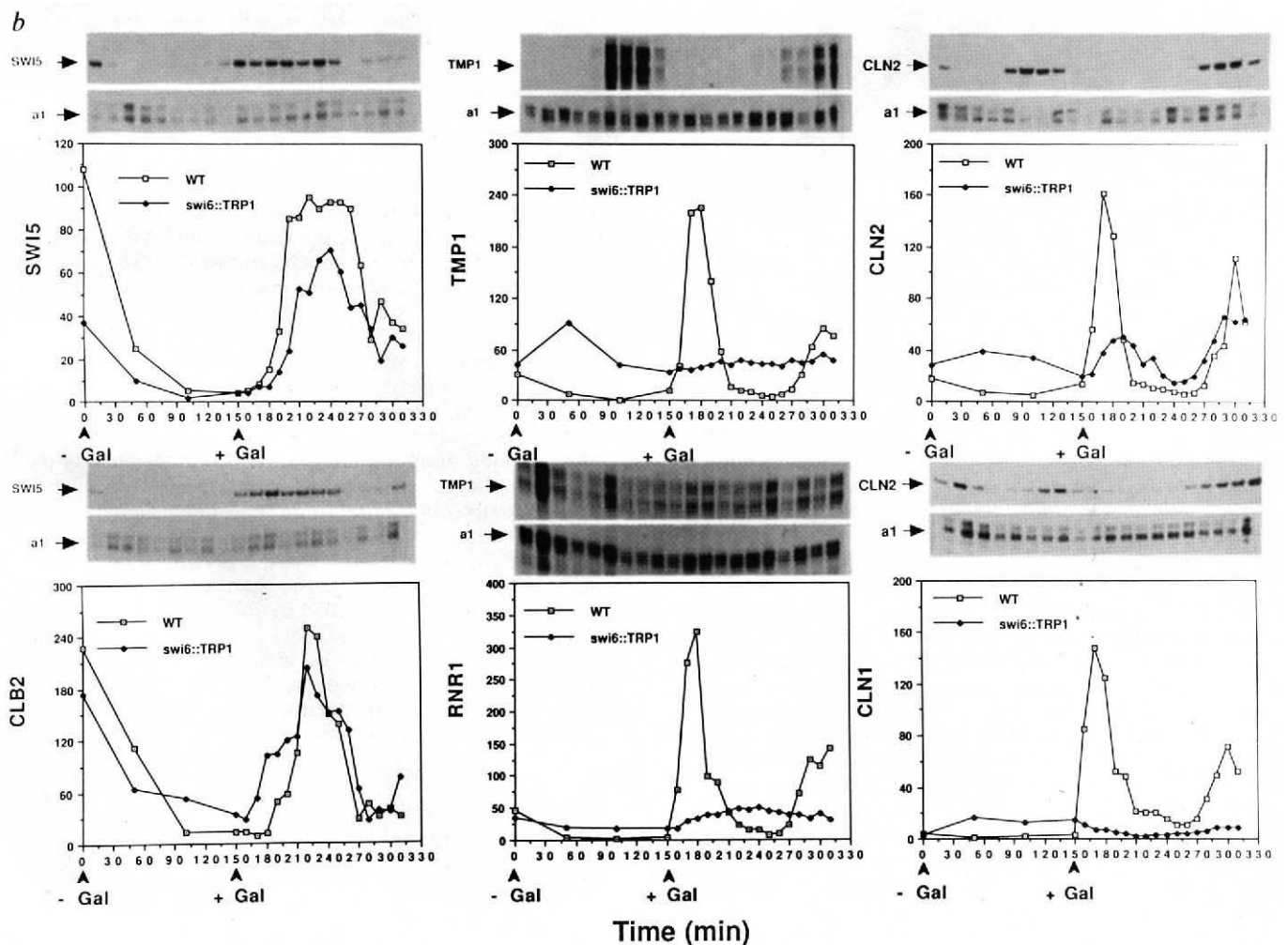


FIG. 2 Footprint and interference analysis of MBF binding to the *TMP1* promoter. *a*, Protection from phenanthroline copper (OP-Cu) cleavage; *b*, methylation interference; *c*, carbethoxylation interference. 'Upper' and 'lower' refer to the strand analysed, and *f* and *b* to free and bound probe respectively. *d*, Summary of interference data. Base pairs conserved in MCBs are in bold typeface. In *a-d*, Greek letters refer to bases whose methylation interferes with binding and numbers refer to bases whose carbethoxylation interferes with binding.

METHODS. *a*, Footprinting using 1.10-phenanthroline copper ion was as described⁹. The MCB-TMP1 oligonucleotide was cloned into the *Sal*I/*Bam*HI sites of pSP72 (Promega). The plasmid (C1991) was cut with *Eco*RI and *Pst*I, the MCB-TMP1 fragment gel-purified and end-labelled with ³²P using T7 DNA polymerase. The probe was mixed (as described for Fig. 1) with MBF partially purified on heparin-agarose and free and bound DNAs separated by gel electrophoresis. The gel was then treated with OP-Cu and both free and retarded (bound) probe were eluted from the gel and precipitated with ethanol. The marker lane (G + A) was a G and A sequencing reaction¹⁰. Equal Cerenkov c.p.m. were loaded onto a denaturing 10% polyacrylamide gel. *b*, Methylation interference was done as described¹¹. The C1991 plasmid was cut with either *Eco*RI or *Hind*III, and end-labelled using [γ -³²P]ATP and T4 polynucleotide kinase. DNA was then partially methylated with 5 μ l dimethylsulphate in 200 μ l cacodylate buffer, pH 7.2, (ref. 10) for 3 min at room temperature and the probe was liberated from the rest of the plasmid by digestion with a second restriction enzyme at the other side of the *TMP1* promoter (*Eco*RI if DNA was labelled at the *Hind*III site and vice versa). The probe was mixed with MBF partially purified on heparin-agarose and free and bound DNA separated by gel electrophoresis. DNA was eluted from the gel, purified by binding on a DEAE membrane, cleaved with piperidine¹⁰ and run on a denaturing 10% polyacrylamide gel. *c*, Carbethoxylation interference¹¹ was as described, with the exception of the DNA modification. The end-labelled probe was denatured by boiling for 3 min and snap-cooling. Diethylpyrocarbonate treatment was for 20 min at 37 °C in 300 μ l cacodylate buffer using 10 μ l diethylpyrocarbonate.

FIG. 3 Effect of mutating *SWI6* on cell-cycle-regulated transcripts. *a*, *TMP1* and *RNR1* RNAs in *swi4* and *swi6* mutants measured by S1 mapping and quantified using a Molecular Dynamics phosphorimager. The ratios of *TMP1* to internal control *MATa1* were 102% for *swi4Δ* and 90% for *swi6Δ* relative to wild type (WT) (100%). The ratios for 'normal' *RNR1* transcripts were 104% for *swi4Δ* and 59% for *swi6Δ*. *b*, Effect of mutating *SWI6* on the cell cycle periodicity of genes containing MCBs and SCBs. Squares, wild type; filled symbols, *swi6::TRP1* (*SWI6* deleted and replaced by the *TRP1* gene). Congenic *SWI6*⁺ and *swi6Δ* strains whose passage through Start depends on expression of *CLN1* from the *GAL1-10* promoter were grown in YEP medium containing 2% raffinose and 2% galactose at 30 °C. At *t*=0 both cultures were collected by filtration, washed in YEP raff medium and reinoculated in YEP raff medium minus galactose (-Gal). After 150 min, galactose was added back to the medium (+Gal). *SWI5*, *TMP1*, *CLN2*, *RNR1* and *CLN1* RNAs were measured by an S1 protection assay²⁶, and *CLB2* RNAs by northern blotting. All RNAs were quantified using a Molecular Dynamics phosphorimager and normalized using a *MATa1* internal control. The ratio of these RNAs to *MATa1* multiplied by 100 is plotted. Autoradiographs of S1 protected fragments due to *SWI5*, *TMP1* and *CLN2* RNAs are shown above (WT) and below (*swi6Δ*) the graphs.

METHODS. *a*, S1 protection assays^{5,26} were done on RNAs extracted from the congenic strains: K1107 (WT), K1899 (*swi4Δ*), and K1970 (*swi6Δ*)¹⁶ growing exponentially in YEPD medium at 30 °C. A 0.71-kb *HindIII*-*BglIII* polymerase chain reaction (PCR) fragment cloned between the *HindIII* and *BamHI* sites of M13mp19 was used as a template for preparing single-stranded *TMP1* probe, which was released by cutting with *HindIII* after Klenow extension. A 0.9-kb *EcoRI*-*BglIII* fragment cloned into M13 mp18 was used for *RNR1*¹⁴. This probe was released with *EcoRI*. *b*, K2771 (*cln1Δ*, *cln2Δ*, *cln3Δ*, YCpGAL-*CLN1*; congenic with K1107) was a meiotic segregant from the diploid K2342 (ref. 16) transformed with YCpGAL-*CLN1*. The *swi6::TRP1* derivative (K2786) was made by transformation. Both strains grow only in the presence of galactose.



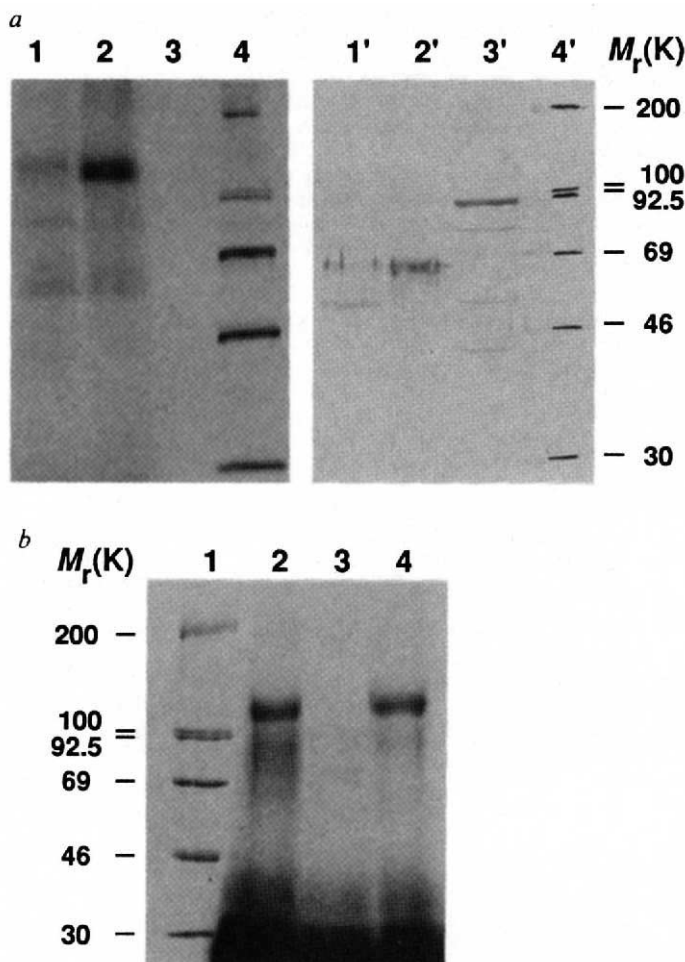


FIG. 4 A 120K protein that is neither Swi4 nor Swi6 can be crosslinked by ultraviolet light to MCBs. *a*, Binding reactions on the partially purified MBF were as described for Fig. 1 in the presence and absence of antibodies against Swi6, except that a bromodeoxyuridine-substituted MCB-TMP1 oligonucleotide was used²¹. After gel electrophoresis, the protein-DNA complexes were covalently crosslinked by ultraviolet irradiation, digested with DNase I, eluted from the gel and covalently bound protein analysed by SDS-PAGE and autoradiography. A single band is detected at ~120K (lane 2) which is strongly reduced if Swi6 antibodies are present during the binding (lane 1). Yeast whole-cell extracts (lane 3) were separated on the same SDS-polyacrylamide gel. An identical crosslinking result was obtained using a crude extract (data not shown). The same gel was rehydrated after autoradiography, proteins were transferred to nitrocellulose and Swi6 assayed using Swi6 antibodies and alkaline phosphatase conjugated to anti-rabbit antibodies. Swi6 is detected not only in the whole-cell extract (lane 3') but also in the crosslinked sample (lane 2': very faint but visible in original prints). In similar gels, Swi6 from the heparin fraction was shown to comigrate with that from crude extracts and Swi4 from both the heparin fraction and from crude extracts and Swi4 from both the heparin fraction and from crude extracts migrated as a 150K protein (data not shown). Molecular weight markers are shown in lanes 4 and 4'. *b*, Binding reactions using the BrdU-substituted oligonucleotide (1×10^6 c.p.m.) were done as in *a*, but instead of using reaction mixtures for gel retardation electrophoresis they were irradiated with ultraviolet light digested with DNase I and analysed by SDS-PAGE and autoradiography. Binding was assayed in the absence of competitor (lane 2) and in the presence of a 50-fold molar excess of cold specific competitor (MCB-TMP1; lane 3). Addition of a 50-fold molar excess of a cold point-mutated MCB-TMP1 (lane 4; 5' and 3' MCBs mutated, see Fig. 1c) does not lead to a significant reduction in the amount of p120 detected. Molecular weight markers are shown in lane 1.

METHODS. For the crosslinking analysis, the 5'-dCCTTAACGCGTCACGG-3' oligonucleotide was annealed to the top strand of the $3 \times$ *Mul* oligonucleotide (Fig. 1 legend), filled up with Klenow in the presence of bromo-dUTP and [α -³²P]dATP, digested with *Bam*HI and gel-purified. Preparative gel retardation assays used $1-2 \times 10^6$ c.p.m. of labelled oligonucleotide and protein-DNA complexes were crosslinked by ultraviolet irradiation (at 302 nm for 12 min). After exposure of the wet gel to X-ray film, the complex was excised from the gel, digested with DNase I and recovered from the gel as described²¹. Protein was precipitated with trichloroacetic acid from the supernatant and analysed by SDS-PAGE and autoradiography.

because it is not part of MBF and because on western blotting it migrates as a 150K protein in the same gel (data not shown). Further, p120 is crosslinked to MCBs by ultraviolet light when partially purified MBF is mixed with BrdU-substituted MCB-TMP1 in solution in the absence but not in the presence of specific competitor (Fig. 4b). In addition, p120 can be co-immunoprecipitated from the reaction with Swi6- but not with Swi4-specific antibodies (data not shown). We therefore suggest that MBF is composed of at least two proteins: Swi6 and a 120K protein that is not Swi4. The detection by silver staining of comparable amounts of swi6 and 120K protein in highly purified preparations of MBF (T.M., unpublished observation) corroborates this. We were not able to crosslink Swi6 to the *TMPI* oligonucleotide, suggesting that Swi6 might not be in intimate contact with the DNA.

Swi6 may play a central part in late G1- or Start-dependent transcription in yeast because it is the common component of two different Start-dependent transcription factors (SBF and MBF). We suggest that Swi6 forms complexes with Swi4 on SCBs but with p120 on MCBs. As Swi4 alone can bind to SCBs *in vitro* and p120 but not Swi6 can be crosslinked to MCBs, we propose that the slightly different DNA sequence specificities of the two types of Swi6-containing complexes are due to the site-specific DNA-binding domains of Swi4 and p120. These domains may prove to be related because the SCB and MCB are similar sequences and can cross-compete in binding assays (Fig. 1a). To explain why the level of *TMPI* transcripts in cycling cells is unaffected by the deletion of *SWI6* whereas extracts from *swi6* mutants fail to form any specific complexes on *TMPI* MCBs *in vitro*, we propose that p120 can still bind MCBs without Swi6, but only with a reduced affinity. Thus it may bind and activate transcription (constitutively) *in vivo* even though it cannot bind *in vitro* when diluted in extracts. This

would be analogous to the partial activation by Swi4 of *CLN2* in the absence of Swi6^{16,22}. Our data indicate that Swi6 is necessary for regulating transcription exerted by Swi4. We suggest by analogy that it has a similar function for p120. *TMPI* RNAs do not drop after galactose removal nor do they increase when it is restored in the *swi6* mutant culture (Fig. 3b), implying that Swi6 acts as a repressor during early G1, G2 and M phases, but as an activator during late G1. We propose that Swi6 be called a modulator, which seems apt for a protein that imposes periodicity on transcription. The finding that the Swi6-related²³ *cdc10* protein is a component of an MBF-like factor in *Schizosaccharomyces pombe*²⁴ suggests that such modulators may exist in many if not all eukaryotes.

The aperiodic transcription observed in *swi6* mutants indicates that cell-cycle changes in the Swi6 protein may impose cell-cycle control. Phosphorylation of swi6 by the CDC28 (or another) kinase could possibly transform it from a repressor to an activator around the time of Start. It is also possible that the changes triggering transcription at Start occur on other components of SBF and MBF (for example, Swi4 and p120) and that Swi6 is required to prevent these proteins adopting an active conformation in the absence of such changes. Our finding of MBF binding activity throughout the cell cycle suggests that changes in an 'activation' function rather than binding *per se* may trigger *TMPI* transcription at Start. □

Received 6 March; accepted 30 April 1992.

- Andrews, B. J. & Herskowitz, I. *J. Biol. Chem.* **265**, 14057-14060 (1990).
- Pizzagalli, A., Valsasini, P., Plevani, P. & Lucchini, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3772-3776 (1988).
- McIntosh, E. M., Atkinson, T., Storms, R. K. & Smith, M. *Molec. cell. Biol.* **11**, 329-337 (1991).
- Gordon, C. B. & Campbell, J. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6058-6062 (1991).
- Nasmyth, K. *Cell* **42**, 225-235 (1985).
- Andrews, B. J. & Herskowitz, I. *Nature* **342**, 803-833 (1989).

7. Taba, M. R., Muroff, I., Lydall, D., Tebb, G. & Nasmyth, K. *Genes Dev.* **5**, 2000–2013 (1991).
8. Ammerer, G. *Genes Dev.* **4**, 299–312 (1990).
9. Kuwabara, M. D. & Sigman, D. S. *Biochemistry* **26**, 7234–7238 (1987).
10. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
11. Sturm, R., Baumruker, T., Franua, B. R. & Herr, W. *Genes Dev.* **1**, 1147–1160 (1987).
12. Guarente, L. & Mason, T. *Cell* **32**, 1279–1286 (1983).
13. Price, C., Nasmyth, K. & Schuster, T. *J. Molec. Biol.* **218**, 543–556 (1991).
14. Elledge, S. J. & Davis, R. W. *Genes Dev.* **4**, 740–751 (1990).
15. Cross, F. & Tinklenberg, A. H. *Cell* **65**, 875–883 (1991).
16. Nasmyth, K. & Dirick, L. *Cell* **66**, 995–1013 (1991).
17. Lydall, D., Ammerer, G. & Nasmyth, K. *Genes Dev.* **5**, 2405–2419 (1991).
18. Surana, U. *et al.* *Cell* **65**, 145–161 (1991).
19. Ghara, J. B. *et al.* *Cell* **65**, 163–174 (1991).
20. Ogas, J., Andrews, B. J. & Herskowitz, I. *Cell* **66**, 1015–1026 (1991).
21. Barberis, A., Widenhorn, K., Vitelli, K. & Busslinger, M. *Genes Dev.* **4**, 849–859 (1990).
22. Moll, T., Dirick, L., Auer, H. & Nasmyth, K. *J. Cell Sci.* (in the press).
23. Breeden, L. & Nasmyth, K. *Nature* **329**, 651–654 (1987).
24. Lowndes, N. F., McInerney, C. J., Johnson, A. L., Fantes, P. A. & Johnston, L. H. *Nature* **355**, 449–453 (1992).
25. Stillman, D. J., Bankier, A. T., Seddon, A., Groenhout, E. G. & Nasmyth, K. *EMBO J.* **7**, 485–495 (1988).
26. Dirick, L. & Nasmyth, K. *Nature* **351**, 754–757 (1991).

ACKNOWLEDGEMENTS. We thank J. Bonkovsky for the MCB-TMP1/CYC-LacZ plasmid and F. Cross for the YCpGal-CLN1 plasmid; G. Ammerer and P. Sorger for heparin-agarose-purified fractions; R. Taba for yeast extracts; H. Tkadletz for preparing the figures; R. Kurzbauer and G. Schaffner for oligonucleotide synthesis and DNA sequencing; and B. Errede, C. Koch, G. Tebb and M. Busslinger for their help and comments on the manuscript.

Functional contacts of a transfer RNA synthetase with 2'-hydroxyl groups in the RNA minor groove

Karin Musier-Forsyth & Paul Schimmel

Department of Biology, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

THE functional analysis of determinants on RNA has been largely limited to molecules that contain naturally occurring ribonucleotides, so little is known about the role of 2'-hydroxyl groups in protein-RNA recognition. A single base pair (G3·U70) in the acceptor stem of tRNA^{Ala} is the principal element for specific recognition by *Escherichia coli* alanine-tRNA synthetase^{1,2}. This tRNA synthetase aminoacylates small RNA helices that contain the G3·U70 base pair. Furthermore, removal of the G3 exocyclic 2-amino group that projects into the minor groove eliminates aminoacylation³. This 2-amino group is flanked on either side by ribose 2'-hydroxyl groups that line the minor groove. Here we use chemical synthesis to construct 32 helices that make deoxy and O-methyl substitutions of individual and multiple 2'-hydroxyl groups near and beyond the G3·U70 base pair and find that functional 2'-hydroxyl contacts are clustered within a few ångströms of the critical 2-amino group. These contacts are highly specific and make a thermodynamically significant contribution to RNA recognition.

Single deoxyribonucleotide substitutions were incorporated into the 3' and 5' strands of the tRNA^{Ala} duplex at the eleven positions shown in Fig. 1 (left). The ability of alanine-tRNA synthetase to aminoacylate each of the eleven variants was compared to aminoacylation of a wild-type (all-ribose) duplex. Initial rates of aminoacylation were determined under conditions of low substrate concentration (substrate concentration $S \ll K_m$). The initial rate was therefore a measure of V_{max}/K_m and was used directly to calculate $\Delta\Delta G^{\ddagger}$, the difference in the free energy of transition-state formation between the wild type and duplex variants⁴. As seen from Fig. 1 and Table 1, single deoxy substitutions at only two positions in the 3' strand (C71 and U70) and just one position in the 5' strand (G4) significantly decreased aminoacylation efficiency ($\Delta\Delta G^{\ddagger} > 1 \text{ kcal mol}^{-1}$).

To determine whether this decrease is due to the removal of specific contacts involving the 2'-OH group acting as a hydrogen-bond donor, single 2'-O-methyl substitutions were made at five positions (Fig. 1, middle). At positions where the deoxy substitution had little effect (C72, C69, G3 and C5), incorporation of a 2'-O-methyl also had little effect on the ability of the enzyme

TABLE 1 The effect of single minor groove substitutions on the free energy of transition state formation

3'-Strand substitutions*	$\Delta\Delta G^{\ddagger}$ ($k_{cal} \text{ mol}^{-1}$) [†]	5'-Strand substitutions§	$\Delta\Delta G^{\ddagger}$ ($k_{cal} \text{ mol}^{-1}$)
dA73	0.0	dG1	0.2
dC72	-0.4	dG2	0.5
dC71	1.8	dG3	0.2
dU70	1.7	dG4	1.2
dC69	0.1	dC5	-0.1
dG68	-0.4	2'-O-MeG3	0.6
2'-O-MeC72	0.7	2'-O-MeC5	0.1
2'-O-MeC71	2.1	dI1	0.3
2'-O-MeC69	0.8	dI2	1.3
		I3	>3.3
		I4	0.8

* Aminoacylation was assayed as described in the legend to Fig. 1 after annealing the 'mutant' 13-base oligonucleotide to a complementary 'wild-type' (all ribose) 8- or 9-base oligonucleotide.

† $\Delta\Delta G^{\ddagger} = RT \ln (V_{max}/K_m)^{mutant} / (V_{max}/K_m)^{wild-type}$. Results are averages of 2–4 experiments with standard deviations of ± 0.0 – $0.4 \text{ kcal mol}^{-1}$, depending on the substrate.

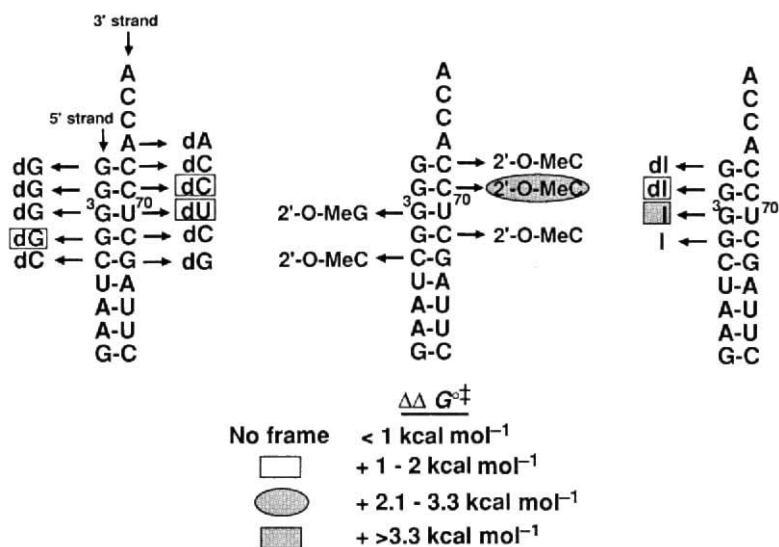
§ Aminoacylation was assayed as described in the legend to Fig. 1 after annealing the 'mutant' 8- or 9-base oligonucleotide to a complementary 'wild-type' (all-ribose) 13-base oligonucleotide.

to aminoacylate the mutant duplex (Table 1). In the case where the deoxy substitution resulted in a significant increase in $\Delta\Delta G^{\ddagger}$ (C71), the 2'-O-methyl substitution gave in an even greater increase (Table 1). This suggests that the 2'-OH of C71 is in close contact with the bound enzyme.

The results of aminoacylation experiments using duplexes containing several deoxyribonucleotide substitutions were consistent with those from single deoxy substitution and illustrated the importance of helix geometry to recognition by alanine-tRNA synthetase. Previous attempts to aminoacylate a 3'-DNA(rA76)/5'-RNA hybrid duplex with alanine were unsuccessful (Fig. 2, duplex a)⁵, but a 3'-RNA/5'-DNA heteroduplex was aminoacylated, albeit at a reduced rate relative to the all-ribose duplex (Fig. 2, duplex f). Our results from single deoxy substitution support the hypothesis that functionally important 2'-OH contacts rather than helical distortion are responsible for the decreased aminoacylation efficiency of the heteroduplexes. A single position in the 5' strand (G4) is sensitive to deoxyribonucleotide incorporation (Fig. 1; Table 1) and this substitution is the major contributor to the $\Delta\Delta G^{\ddagger}$ value of the 3'-RNA/5'-DNA duplex (Fig. 2, duplex f). On the other hand, the functional importance of 2'-hydroxyls at several positions of the 3' strand (U70, C71 in Table 1, and C75, A76 (data not shown)) explains, at least in part, why aminoacylation of the 3'-DNA(rA76)/5'-RNA duplex is not detected. The effects of triple (dU70, dC71 and dC75) and double (dC71 and dC75) deoxy substitutions confirm this hypothesis, because these primarily ribose duplexes should have an A-like helix configuration but are not aminoacylated by alanine-tRNA synthetase (Fig. 2, duplexes c and d).

The relative importance of specific 2'-hydroxyls versus helix conformation was also explored by preparing 'minimum-ribose' substrates. These substrates contain ribonucleotides only at those locations where the 2'-hydroxyl was important in the single-deoxy substitution experiments (Table 1, and data not shown). A minimum-ribose 3' strand, annealed to an all-ribose 5' strand (Fig. 2, duplex e), showed little decrease in aminoacylation efficiency relative to the wild-type duplex. Similarly, an all-ribose 3' strand annealed to d(GG)GGd(CUAAG) or to a minimum-ribose 5' strand, was efficiently aminoacylated by alanine-tRNA synthetase (Fig. 2, duplexes g and h). Surprisingly, the duplex generated by annealing 3'- and 5'-strand minimum-ribose oligonucleotides could still be aminoacylated but was a poorer substrate for the synthetase (Fig. 2, duplex k).

FIG. 1 Duplex variants containing single minor groove substitutions. Oligonucleotides were prepared by automated solid phase chemical synthesis on a Gene Assembler Plus (Pharmacia) synthesizer using the phosphoramidite method¹³. Each duplex consists of a 13-nucleotide 3' strand annealed to an 8- or 9-nucleotide 5' strand; arrows point to the single base substitutions. Fully protected phosphoramidite monomers were from ChemGenes (Waltham, MA) or Glen Research (Sterling, VA). Extinction coefficients at 260 nm were estimated to be $10.7 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ for the 13-base oligonucleotide; $8.9 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$, 9-base oligonucleotide; $7.8 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$, 8-base oligonucleotide⁵. Results of assays comparing the aminoacylation efficiency of the wild-type (all-ribose) duplex with each of the 20 duplex variants are indicated by the absence ($\Delta\Delta G^{\ddagger} < 1 \text{ kcal mol}^{-1}$) or presence ($\Delta\Delta G^{\ddagger} > 1 \text{ kcal mol}^{-1}$) of a symbol around the base change. Values of $\Delta\Delta G^{\ddagger}$ were calculated as described in Table 1 legend. Aminoacylation was assayed^{14,15} at 25 °C with 0.5 μM purified *E. coli* alanine-tRNA synthetase in a reaction mixture containing 50 mM HEPES (pH 7.5), 20 mM KCl, 10 mM MgCl_2 , bovine serum albumin at 0.1 mg ml^{-1} , 20 mM 2-mercaptoethanol, 22 μM [^3H]alanine, 4 mM ATP and 10 μM annealed duplex. For initial rate determinations, 18- μl aliquots were removed from this reaction mixture and spotted onto Whatman 3MM filter pads at several time points out to 6 min. Duplexes were annealed before each assay by heating complementary strands to 80 °C for 3 min in



With only three ribonucleotides in the double-stranded region and no ribose-ribose base pairs, it is unlikely that this duplex maintains an A-form geometry. An altered helix configuration of duplex *k* may prevent enzyme access to critical functional groups and contribute to the observed $\Delta\Delta G^{\ddagger}$ of 1.8, which is $\sim 1 \text{ kcal mol}^{-1}$ greater than would be expected if the multiple deoxy substitutions were strictly additive ($\Delta\Delta G^{\ddagger}$ of duplex *e* plus $\Delta\Delta G^{\ddagger}$ of duplex *h* is 0.8 kcal mol^{-1}). Incorporation of one additional deoxy substitution at position G4 of the minimum-ribose duplex eliminates detectable aminoacylation (Fig. 2, duplex *l*).

Several other hybrid duplexes were synthesized and tested for aminoacylation. For example, an all-ribose 5' strand annealed to d(CUUAG)CUD(CCACC)A was not aminoacylated by alanine-tRNA synthetase (Fig. 2, duplex *b*). This is the result expected on the basis of the reduced aminoacylation observed when single deoxy substitutions are incorporated at C71 ($\Delta\Delta G^{\ddagger} = 1.8$; Table 1) and C75 ($\Delta\Delta G^{\ddagger} = 1.6$; data not shown). No aminoacylation was detected either when this same 3' strand was annealed to d(GG)GGd(CUAAG) (Fig. 2, duplex *j*). Finally, a duplex containing a double-deoxy substitution at the G3-U70 base pair is poorly aminoacylated, as would be expected from the single deoxy substitution experiments (Fig. 2, duplex *i*).

The role of 2'-hydroxyls in defining RNA structure and stability is well established⁶⁻⁸. It has been shown that the 2'-hydroxyl groups of RNA mediate the association between ribozymes and their RNA substrates and are involved in both the substrate binding and cleavage steps⁹⁻¹². Here the mediation of an RNA-protein interaction by the 2'-hydroxyls on the RNA is summarized by a model of the tRNA^{Ala} acceptor helix shown in Fig. 3. Removal of the atoms shown in pale blue had little or no effect on the aminoacylation efficiency of the duplex. Of all the single substitutions examined, incorporation of inosine at position 3 (which removes the 2-amino group shown in red) had the greatest effect on helix recognition by alanine-tRNA synthetase, eliminating detectable aminoacylation. Other functionally important atoms identified here are shown in white (Fig. 3): the 2'-hydroxyls of G4, U70 and C71 are clustered within $\sim 5 \text{ \AA}$ of the exocyclic 2-amino group of G3. Experiments using 2'-O-methyl-substituted nucleotides indicate that unimportant 2'-OH groups are not in close contact with the enzyme so functional 2'-hydroxyls may be involved in hydrogen bonding with the synthetase either directly or through bridging water molecules.

20 mM sodium phosphate (pH 7.0) and then cooling to 4 °C. Results were scaled by a factor of 1.5 to correct for incomplete precipitation of the 13-base oligonucleotide onto the Whatman filter pads used in the assays⁵.

The tight clustering of minor groove 2'-OH contacts proximal to the G3 2-NH₂ group invites analysis of this area for other strategic interactions. We previously established that a G4 substitution by inosine had little effect on aminoacylation³. Duplexes containing single deoxyinosine substitutions at positions 1 and 2 (dI1 and dI2) of the 5' strand were also tested for their ability

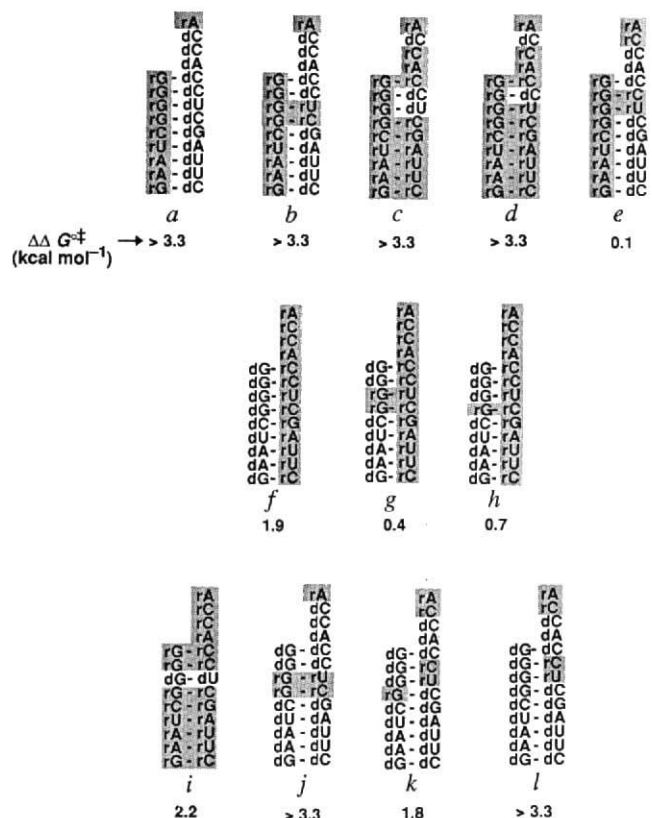


FIG. 2 Duplex variants containing multiple minor groove deoxynucleotide substitutions. Duplexes contain 3'-strand substitutions (a-e), 5'-strand substitutions (f-h), and both 3'- and 5'-strand substitutions (i-l). Ribonucleotides are shaded and values of $\Delta\Delta G^{\ddagger}$ were calculated as described in Table 1 legend.

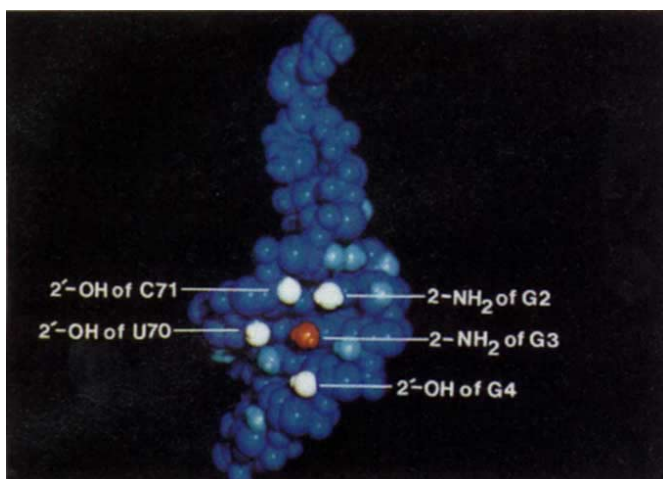


FIG. 3 Space-filling model showing a minor groove view of the $tRNA^{Ala}$ acceptor helix. Coordinates for $tRNA^{Phe}$ were used to generate this model, which shows the first 7 base pairs and the 3'-ACCA end (top) of the $tRNA^{Ala}$ acceptor helix. Dark blue, atoms unchanged by the single-base substitution experiments; light blue, functional groups (-OH and -NH₂) whose removal (substitution with H) has little effect ($\Delta\Delta G^{\ddagger} < 1 \text{ kcal mol}^{-1}$) on aminoacylation efficiency; white, functional groups whose removal has a significant effect ($\Delta\Delta G^{\ddagger} > 1 \text{ kcal mol}^{-1}$) on aminoacylation efficiency; red, functional group (-NH₂) whose removal eliminated detectable aminoacylation by alanine-tRNA synthetase ($\Delta\Delta G^{\ddagger} > 3.3 \text{ kcal mol}^{-1}$).

to be aminoacylated by alanine-tRNA synthetase (Fig. 1, right). The dI1-substituted duplex was efficiently aminoacylated but the dI2 substitution resulted in a significant increase in $\Delta\Delta G^{\ddagger}$ (Table 1). This effect can probably be attributed to the removal of the exocyclic 2-amino group, because the duplex containing deoxyguanosine at position 2 was efficiently aminoacylated (Table 1). The G2 2-NH₂ group falls within the radius of functional 2'-OH contacts (Fig. 3). These experiments suggest that the 2-NH₂ group of G2 plays a role in synthetase recognition, but that it is not as important as the 2-NH₂ group of G3 (Table 1). This agrees with experiments showing that the transfer of G3·U70 into RNA substrates with a C2·G71 (ref. 1) or a U2·A71 (ref. 4) base pair allows aminoacylation with alanine, albeit at a reduced rate compared with the G3·U70, G2·C71 substrate. The contribution of the individual C71 or U70 2'-OH groups to the free energy of transition state formation is greater than that of the G2 2-NH₂ group. This comparison shows that, within the cluster of minor groove interactions flanking the 2-NH₂ group of G3, functional contacts with 2'-OH groups are quantitatively at least as important as base-specific interactions. □

Novel form of growth cone motility involving site-directed actin filament assembly

Paul Forscher*, Chi Hung Lin* & Corey Thompson†

* Department of Biology, Yale University, New Haven, Connecticut 06511, USA

† Department of Cell Biology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

REGULATION of cytoskeletal structure and motility by extracellular signals is essential for all directed forms of cell movement and underlies the developmental process of axonal guidance in neuronal growth cones. Interaction with polycationic microbeads can trigger morphogenic changes in neurons and muscle cells normally associated with formation of pre- and postsynaptic specializations^{1,2}. Furthermore, when various types of microscopic particles are applied to the lamellar surface of a neuronal growth cone or motile cell they often exhibit retrograde movement at rates of 1–6 $\mu\text{m min}^{-1}$ (refs 3–6). There is strong evidence that this form of particle movement results from translocation of membrane proteins associated with cortical F-actin networks, not from bulk retrograde lipid flow^{4,5,7} and may be a mechanism behind processes such as cell locomotion, growth cone migration and capping of cell-surface antigens^{6,8,9}. Here we report a new form of motility stimulated by polycationic bead interactions with the growth-cone membrane surface. Bead binding rapidly induces intracellular actin filament assembly, coincident with a production of force sufficient to drive bead movements. These extracellular bead movements resemble intracellular movements of bacterial parasites known to redirect host cell F-actin assembly for propulsion. Our results suggest that site-directed actin filament assembly may be a widespread cellular mechanism for generating force at membrane-cytoskeletal interfaces.

When polycationic microbeads were applied to the surface of *Aplysia* bag cell growth cones maintained in serum-free medium, two distinct modes of membrane-bound bead movement were observed: (1) linear retrograde bead movements resulting from coupling to submembranous F-actin flow as previously described⁴ (referred to here as 'flow-coupled' transport) and (2) a novel form of bead movement characterized by formation of distinct comet tail-like structures behind beads moving on the extracellular membrane surface. We refer to these structures as 'inductopodia' because they seem to be induced by bead-membrane interactions. We used video-enhanced differential interference contrast (DIC) and fluorescence digital imaging techniques to investigate the mechanism of inductopodia formation. Figure 1 contrasts the two modes of movement described above. In Fig. 1a, two membrane-bound beads are located in the area of interest: the uppermost bead is undergoing flow-coupled transport; the lower bead is forming an inductopodium. Dynamic aspects of these two processes are shown in the Fig. 1b. Bead A undergoes typical flow-coupled (rearward) transport throughout the sequence; by contrast, bead B triggers inductopodia formation immediately after membrane binding. This inductopodium grew in an anterograde direction and resulted in net bead-translocation against the F-actin flow for about 75 s. Note that during the period of net anterograde bead movement (15–90 s), the inductopodium tail forming behind the bead continues retrograde movement, that is, the tail appears to be coupled to the F-actin flow. At the 90 s time point, the induced component of bead movement spontaneously stops and the beads now move solely by flow-coupled transport along with the residual inductopodium structure for the remainder of the recording. Bead displacement is plotted as a function of time for both beads in Fig. 1c. Note that the flow-coupled bead (A) moved in a retrograde (positive) direction throughout the period

Received 27 February; accepted 8 May 1992.

- Hou, Y.-M. & Schimmel, P. *Nature* **333**, 140–145 (1988).
- McClain, W. H. & Foss, K. *Science* **240**, 793–796 (1988).
- Musier-Forsyth, K. *et al. Science* **253**, 784–786 (1991).
- Franklyn, C., Shi, J.-P. & Schimmel, P. *Science* **255**, 1121–1125 (1992).
- Musier-Forsyth, K., Scaringe, S., Usman, N. & Schimmel, P. *Proc. natn. Acad. Sci. U.S.A.* **88**, 209–213 (1991).
- Wang, A. H.-J. *et al. Nature* **299**, 601–604 (1982).
- Dock-Bregeon, A. C. *et al. Nature* **335**, 375–378 (1988).
- Hall, K. B. & McLaughlin, L. W. *Biochemistry* **30**, 10606–10613 (1991).
- Bevilacqua, P. C. & Turner, D. H. *Biochemistry* **30**, 10632–10640 (1991).
- Chowrira, B. M. & Burke, J. M. *Biochemistry* **30**, 8518–8522 (1991).
- Perreault, J.-P., Labuda, D., Usman, N., Yang, J.-H. & Cedergren, R. *Biochemistry* **30**, 4020–4025 (1991).
- Pyle, A. M. & Cech, T. R. *Nature* **350**, 628–631 (1991).
- Scaringe, S. A., Franklyn, C. & Usman, N. *Nucleic Acids Res.* **18**, 5433–5441 (1990).
- Schreier, A. A. & Schimmel, P. *Biochemistry* **11**, 1582–1589 (1972).
- Jasin, M., Regan, L. & Schimmel, P. *J. biol. Chem.* **260**, 2226–2230 (1985).

ACKNOWLEDGEMENTS. We thank D. Buecher and A. Shepard for comments on the manuscript. This work was supported by a grant from the NIH.

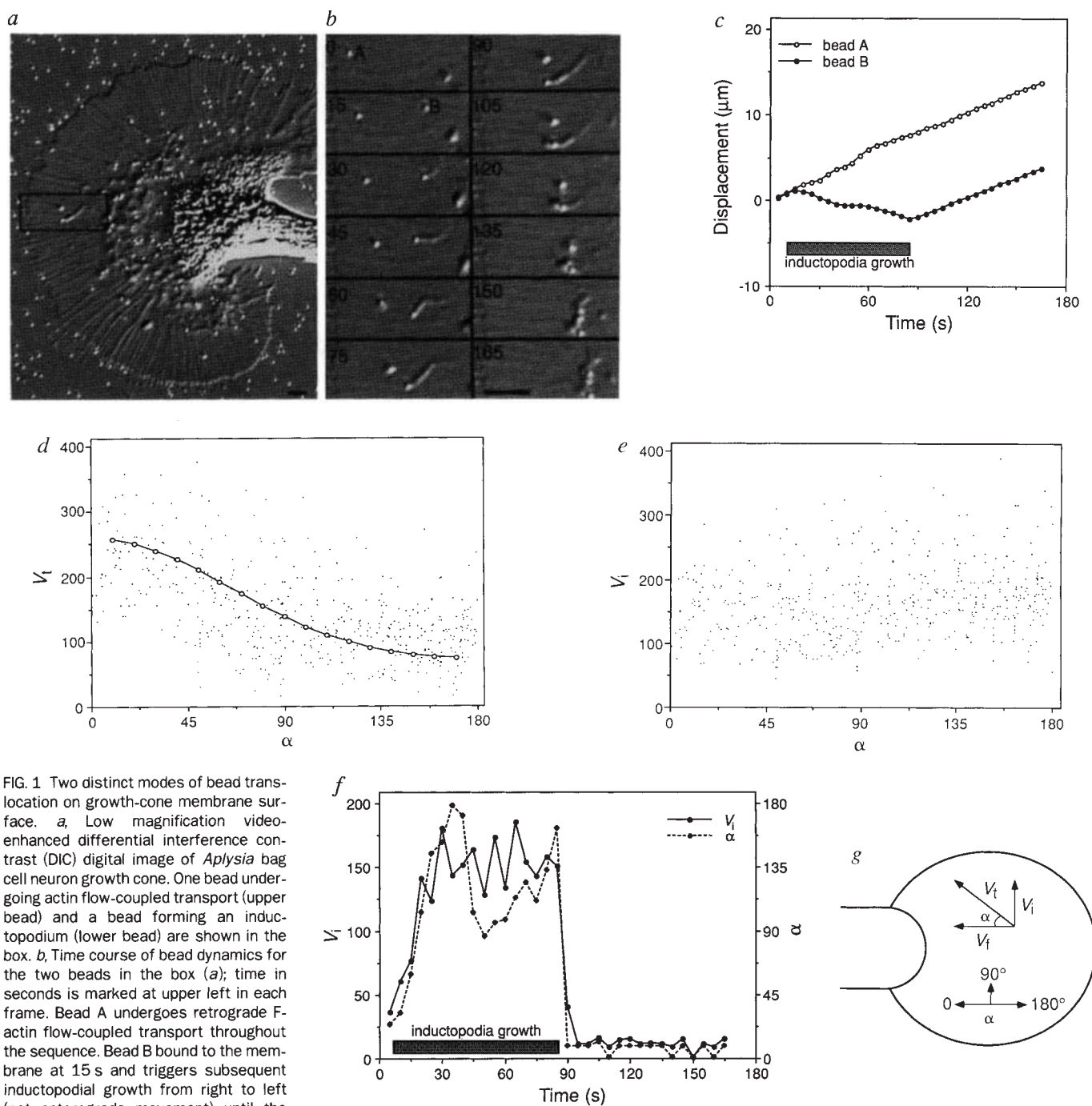


FIG. 1 Two distinct modes of bead translocation on growth-cone membrane surface. *a*, Low magnification video-enhanced differential interference contrast (DIC) digital image of *Aplysia* bag cell neuron growth cone. One bead undergoing actin flow-coupled transport (upper bead) and a bead forming an inductopodium (lower bead) are shown in the box. *b*, Time course of bead dynamics for the two beads in the box (*a*); time in seconds is marked at upper left in each frame. Bead A undergoes retrograde F-actin flow-coupled transport throughout the sequence. Bead B bound to the membrane at 15 s and triggers subsequent inductopodial growth from right to left (net anterograde movement) until the 90 s time point. From 90 to 165 s, both beads move from left to right through flow-coupled transport. Note that inductopodial structure was not stable over time: recently assembled structures near bead-binding sites are always more distinct than older structures distant from beads. *c*, Total bead displacement as a function of time for beads A ($\circ$) and B ($\bullet$). Movements are relative to a growth cone axis placed parallel to the F-actin flow with retrograde movement taken to be positive. Note that total displacement of bead B is anterograde during the period of inductopodial growth. *d*, Total bead velocity (V_t) as a function of translocation angle (α , see *g*) relative to the F-actin flow for beads associated with inductopodia (observations from 70 beads, 10 growth cones). V_t decreases as α approaches 180° (against the flow), as expected if the flow is contributing an angle-dependent velocity component to net bead rates. The points ($\circ$) were generated by a theoretical model for predicting V_t as a function of α if V_t and V_f are simply additive. *e*, Flow-corrected velocity plotted as a function of translocation angle, α . *f*, V_t and α plotted as functions of time for bead B. The period of inductopodial growth is indicated by the shaded bar. All scale bars, $5 \mu\text{m}$.

METHODS. *Aplysia* bag cell neurons were maintained *in vitro* under serum-free conditions. Artificial seawater solutions were prepared as previously

described¹². A custom-made laminar flow cell was used for cell perfusion that enabled rapid exchange of external solutions¹³. Polystyrene microbeads (240 nm diameter; Polysciences) were derivatized with polyethyleneimine as previously described⁴. All experiments were done at room temperature. Optics and digital image processing during primary data acquisition were similar to those described in ref. 13. Bead dynamics were analysed as follows: experiment image records stored on a laser disk recorder were redigitized and bead x, y coordinates determined at 5-s sampling intervals to calculate bead velocities. Inductopodia velocities (V_i) were deduced by vector analysis: for each time interval, a total observed bead velocity vector was constructed, the measured retrograde F-actin flow vector component was then subtracted from V_t to obtain inductopodia velocity (V_i , *g*). The angle between the observed bead movement and the direction of F-actin flow (α), between V_t and V_f , was calculated using standard trigonometry. We developed software for an Imaging Technologies (Bedford, Mass.) Series-151 image processor for determining bead positions and measuring velocities. Bead position and velocity data was ported into a spread sheet program (Microsoft Excel) for analysis of bead population behaviour, F-actin flow correction and calculation of α values.

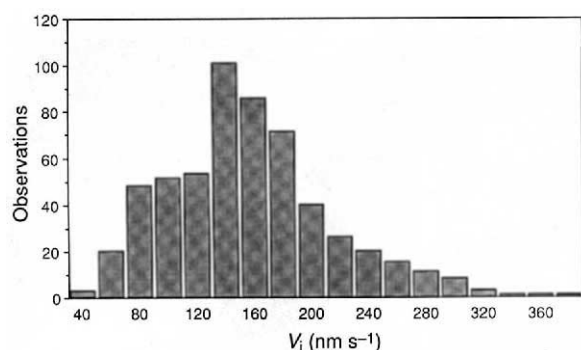


FIG. 2 Histogram of observed F-actin flow-corrected bead velocities obtained during active inductopodial growth. The average V_i was $166 \pm 2.4 \text{ nm s}^{-1}$ ($\pm$ s.e.m., 10 growth cones, 70 beads, >500 observations). The average F-actin flow rate (V_f) was $92.8 \pm 6.1 \text{ nm s}^{-1}$.

of observation; whereas bead B exhibited net anterograde (negative) displacement during the period of inductopodial growth (15–90 s).

Several observations suggested that beads associated with inductopodia move at a total rate (V_t) that is the sum of a constant component contributed by the retrograde F-actin flow (V_f) and a separate bead-induced velocity component (V_i): (1) bead velocities measured from an external frame of reference varied systematically as a function of angle (α) relative to the retrograde F-actin flow. Highest velocities were associated with beads moving parallel to and in the same direction as the retrograde F-actin flow (Fig. 1d). (2) Beads with curvilinear paths always appeared to accelerate as their trajectories became aligned with the flow. (3) Treatment of polycationic beads with a nonspecific blocking protein (1 mg ml^{-1} bovine serum albumin or chicken egg albumen) reversibly inhibited inductopodia formation but did not block flow-coupled bead movements, indicating that the two phenomena were additive. In addition, negatively charged beads were never observed to stimulate inductopodia formation, although they sometimes adhered to growth cone membranes and underwent flow-coupled transport. These results suggested that a critical positive charge density was

necessary on the bead surface to trigger inductopodia formation and that functional separation of the two velocity components (V_i and V_f) could be achieved by experimental manipulation of bead surface charge.

To analyse bead movements associated with inductopodia quantitatively, we used vector analysis of digitized bead tracks to calculate V_i from measured values of V_t and V_f (Fig. 1, g). It was then possible to analyse the bead-induced component of movement, V_i , as a function of time and translocation angle, α . After subtracting V_f from total bead velocities, V_i seemed to be relatively insensitive to α (Fig. 1e). Beads associated with inductopodia moved along both linear and curvilinear paths and with no apparent preference for α , that is, bead translocation trajectories seemed to be isotropic. Figure 1f is a plot of V_i and α over time for bead B during and after inductopodium formation. Note that during the inductopodial growth interval, V_i remained relatively constant despite a full 90° shift in α . A histogram of observed V_i values is shown in Fig. 2. The average V_i was $166 \pm 2.4 \text{ nm s}^{-1}$.

To investigate the structural basis of inductopodia, growth cones were chemically fixed and permeabilized under direct observation by video camera and F-actin distributions were visualized using rhodamine–phalloidin. Inductopodia colocalized with bright comet-shaped tails of F-actin (Fig. 3). F-actin tails were specifically associated with inductopodia formation: beads undergoing flow-coupled transport did not cause measurable alterations in underlying F-actin structure (Fig. 3, solid white arrowheads). These observations strongly suggested that inductopodia formation was due to bead-induced actin filament assembly.

To test this hypothesis, cells were exposed to pulses of cytochalasin B to inhibit actin filament assembly. Figure 4 illustrates the effects of pulses of cytochalasin on inductopodia-related bead movement. The acute effects of blocking F-actin assembly are shown in Fig. 4a: after 15 s in cytochalasin B, inductopodium (bead B) elongation is totally inhibited. Note that retrograde movement of residual F-actin networks persists during the initial phase of cytochalasin treatment, despite rapid inhibition of actin filament assembly as previously described^{4,13}. More importantly, the inductopodium moves centripetally without further tail elongation, in synchrony with both retracting F-actin and membrane-bound beads (arrow and bead A).

FIG. 3 Inductopodia tails colocalize with F-actin. *a*, DIC image of two inductopodia moving roughly perpendicular to the actin flow. Open arrowheads point to beads at the inductopodia heads. Solid white arrowhead points to a bead undergoing F-actin flow-coupled transport. Note lack of observable lamellar perturbation associated with flow-coupled bead. *b*, Filamentous actin distribution visualized with rhodamine-conjugated phalloidin in the same field as *a*. Note that there is no observable alteration of F-actin structure associated with bead undergoing flow-coupled transport (solid white arrowhead marks bead position). Scale bar, $5 \mu\text{m}$. Rhodamine–phalloidin staining of F-actin as in ref. 13.

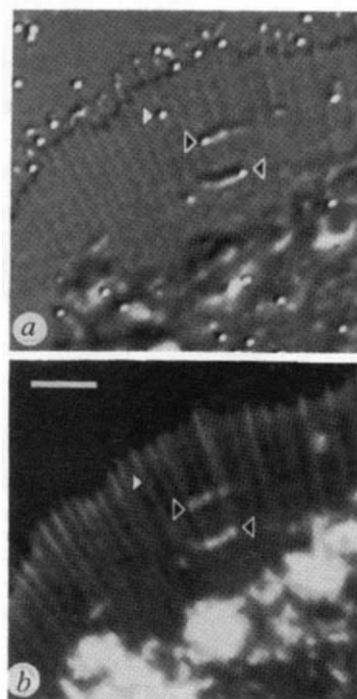


FIG. 4 Cytochalasin B treatment rapidly inhibits inductopodial growth. *a*, Time course of initial effects of cytochalasin B treatment. Timepoints are indicated at upper left before (negative timepoints) and during (CB 0–105) cytochalasin B treatment. Bead A is flow-coupled throughout this sequence; whereas, bead B triggers rapid inductopodial growth accompanied by lateral bead translocation. During cytochalasin B treatment (CB 0–105) lateral bead movement stops completely and the entire inductopodial structure moves a unit in a retrograde direction with the retracting actin network (○ in *c*). White arrow marks successive positions of a bead being transported centripetally by mechanical coupling to underlying retracting F-actin (compare with refs 4, 13). Note that this bead remains a fixed distance from the residual inductopodial structure as both undergo retrograde transport, suggesting that the retracting actin network is behaving like a rigid gel and not a fluid structure. *b*, Examples of inductopodial recovery after cytochalasin treatment from the same growth cone. White arrowheads indicate an inductopodium under control conditions (negative times), during 60 s of cytochalasin B treatment and during 60 s of recovery (Rec 15–60). Inductopodium formation resumed during recovery (● in *c*). Black arrowheads indicate a flow-coupled bead that triggers an inductopodium immediately on cytochalasin B washout. *c*, V_i as a function of time showing inhibition of growth during cytochalasin B treatment (○, *a*, bead B) and recovery growth during drug washout (●, *b*, white arrowheads). Scale bar, 5 μm .

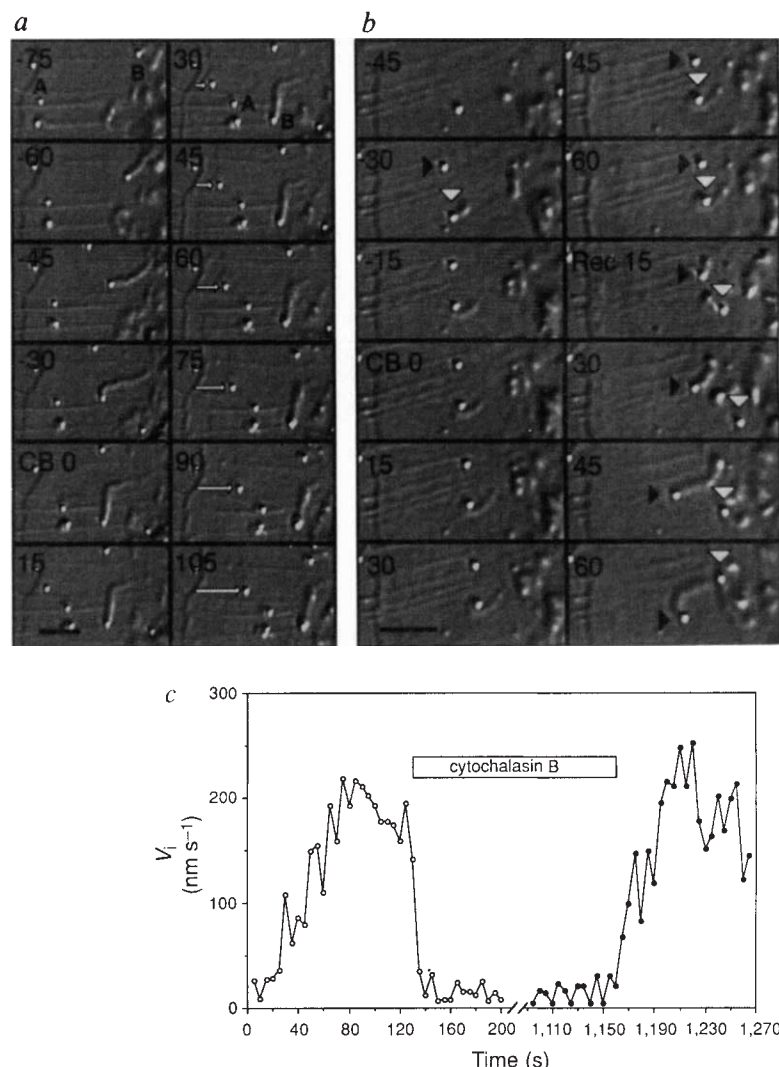


Figure 4*b* illustrates inductopodia dynamics during recovery from another cytochalasin treatment in the same growth cone. On washout of the drug (Fig. 4*b*, Rec 15) inductopodia often immediately resumed growth (white arrowheads) accompanied by bead translocation at control rates. *De novo* formation of inductopodia coinciding with relief from cytochalasin block was also observed (black arrowheads). As expected, V_i dropped rapidly during cytochalasin treatment and recovered rapidly during drug washout (Fig. 4*c*).

These results are, to our knowledge, the first direct demonstration of actin filament assembly stimulated by binding of an extracellular substrate to growth cone lamellae. We have shown that actin assembly can be tightly coupled to the production of forces capable of driving lateral bead movements on membrane surfaces during inductopodia formation. These forces clearly could be involved in directed transport of membrane proteins and/or lipids involved in stimulus-evoked structural remodelling and cell migration. The striking similarity of the comet-shaped actin tails described here to F-actin structures formed during intracellular movements of the bacterial pathogen *Listeria monocytogenes* may be relevant to the mechanism of bead movement in growth cones described here. After entering a host cell, actin filament assembly coupled to filament capping and cross-linking occurs at actin nucleation sites localized to one end of the *Listeria*^{10,11}. This chain of events seems to be sufficient to drive this bacterial parasite rapidly through the host cell's cytoplasm without participation of conventional myosin¹⁰. Polycationic bead translocation on the growth cone surface could

occur by a similar mechanism involving sequestration of endogenous actin nucleating proteins by fixed positive charges on bead surfaces. Inductopodia sensitivity to preadsorption of beads with nonspecific blocking proteins suggests that beads must be able to recruit a critical density of membrane proteins to induce actin filament assembly. Rapid, reversible inhibition of inductopodia formation by cytochalasin B also suggests that actin assembly has a key role in bead translocation, although participation of a myosin-like motor has not yet been ruled out. These interactions mimic membrane-cytoskeletal remodelling associated with target recognition and/or synaptogenesis and may provide a model for addressing the molecular basis of these developmental events. □

Received 27 February; accepted 29 April 1992.

- Peng, H. B., Cheng, P. & Luther, P. W. *Nature* **292**, 831–834 (1981).
- Peng, H. B., Markey, D. R., Muhlach, W. L. & Pollack, E. D. *Synapse* **1**, 10–19 (1987).
- Bray, D. *Proc. natn. Acad. Sci. U.S.A.* **65**, 905–910 (1970).
- Forscher, P. & Smith, S. J. *Optical Microscopy for Biology* 459–471 (Wiley-Liss, New York, 1990).
- Sheetz, M. P., Turney, S., Qian, H. & Elson, E. L. *Nature* **340**, 284–288 (1989).
- Bray, D. & White, J. G. *Science* **239**, 883–888 (1988).
- Lee, J., Gustafsson, M., Magnusson, K. & Jacobson, K. *Science* **247**, 1229–1233 (1990).
- Mitchison, T. J. & Kirschner, M. *Neuron* **1**, 761–772 (1988).
- Smith, S. J. *Science* **242**, 708–715 (1988).
- Dabiri, G. A., Sanger, J. M., Portnoy, D. A. & Southwick, F. S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6068–6072 (1990).
- Tilney, L. G., Connelly, P. S. & Portnoy, D. A. *J. Cell Biol.* **111**, 2979–2988 (1990).
- Forscher, P. & Smith, S. J. *J. Neurosci.* **7**, 3600–3611 (1987).
- Forscher, P. & Smith, S. J. *J. Cell Biol.* **107**, 1505–1516 (1988).

ACKNOWLEDGEMENTS. This work was supported by grants from the NIH and the McKnight Endowment Fund for Neuroscience to P.F.

Gene manipulation on plastic plates

Masato Mitsuhashi, Cylia Keller and Tatsuo Akitaya

Oligonucleotide-immobilized plastic plates provide a rapid and easy-to-use tool for use in mRNA research, including the long-term storage, amplification and detection of the specific message, *in vitro* synthesis of both strands of sense and antisense mRNA, and ligation of cDNA to other DNA molecules.

MESSENGER RNA (mRNA) provides functional genetic information extracted from genomic DNA. Its expression varies among different cells and tissues during the cellular life cycle. However, the mechanisms by which cells express the right mRNA at the right time are still very much a mystery. The main obstacle has been the difficulty in purifying useful amounts of mRNA, which is due in part to the high susceptibility of mRNA to degradation by ribonucleases.

Rapid and efficient purification procedures for mRNA have been developed by various laboratories¹. Furthermore, to identify specific mRNA molecules, procedures such as northern blot analysis², RNase protection assays³, and polymerase chain reaction⁴, have been established. However, mRNA research still involves time-consuming procedures and close attention must be paid to the problems of RNA degradation.

Here, we introduce a multipurpose interface system for research involving mRNA, which can be used in a wide variety of applications by persons without any prior training in molecular biology.

Applications

Figure 1a illustrates the immobilization of oligonucleotides onto plastic microtitre plates. As the oligonucleotide contains single-stranded (ss) poly(dT) sequences, only poly(A)⁺ mRNA is captured (hybridized) onto the plate when lysates of cells or tissues are applied to the plate (Fig. 1b). After unbound non-mRNA is removed by simple aspiration, cDNA can be synthesized using the trapped mRNA as a template (Fig. 1c). One advantage of this system is that fresh buffer containing reverse transcriptase can be replaced many times in order to synthesize full-length cDNA. Furthermore, because synthesized cDNA is much more stable than mRNA, the genetic information of the trapped mRNA can be stored in the form of cDNA for long periods of time. To go from starting cellular materials to cDNA using this system ordinarily takes about three hours.

Once the first strand of cDNA is

synthesized on the plate, it is also possible to synthesize double-stranded (ds) cDNA by adding RNaseH, ligase, DNA polymerase and appropriate buffer (Fig. 1d). Because only the antisense strand of cDNA is covalently immobilized onto the plate at its 5' end, it is possible to separate sense strand cDNA in solution and antisense cDNA on the plate by a simple denaturing procedure.

Sense cDNA can be used for various amplification procedures with specific primers (Fig. 1e), whereas antisense cDNA on the plate can be used for hybridization to the specific probe, as shown in dot blot analysis⁵. However, in contrast to conven-

tional dot blot analysis, in which a large fraction of nucleotides in RNA/DNA is used for crosslinking to nylon membrane, but not for hybridization to the probe, all the nucleotides on the plate are available for hybridization. Furthermore, sense cDNA can be repeatedly synthesized from immobilized antisense cDNA (Fig. 1e-d).

Another advantage of this system is that it is possible to synthesize antisense mRNA from the plate by simply adding RNA polymerase, ribonucleotides and appropriate buffer, as the immobilized oligonucleotide can be provided with specific RNA promoter sequences upstream of the poly(dT) sequences (Fig. 1f). Furthermore, after ligation of double-stranded adaptor at the free end of ds-cDNA on the plate, as shown in Fig. 1g, ligated multiple adaptors can be removed by digestion with restriction enzyme, followed by aspiration of the reaction mixture. This eliminates the need for column chromatography as recommended by other protocols⁶. Once double-stranded adaptor is ligated, sense mRNA can be synthesized by adding the appropriate RNA polymerase for the promoter provided, ribonucleotides and appropriate buffer (Fig. 1h). In other words, if a small quantity of valuable mRNA is treated on the plate, either strand of its cDNA or mRNA can be reproduced and amplified many times.

Finally, synthesized ds-cDNA on the plate can be removed by restriction enzyme digestion and ligated to other DNA molecules because the immobilized oligonucleotide contains multiple restriction enzyme sites upstream of RNA promoter sequences (Fig. 1f).

In summary

The new assay system described for mRNA is designed for use not only by experts in molecular biology but also for researchers who would like to apply advanced techniques in molecular biology to their own fields of research. The method is intended to be easy to use and does not require prior purification of RNA or mRNA. In addition, short processing times and the use

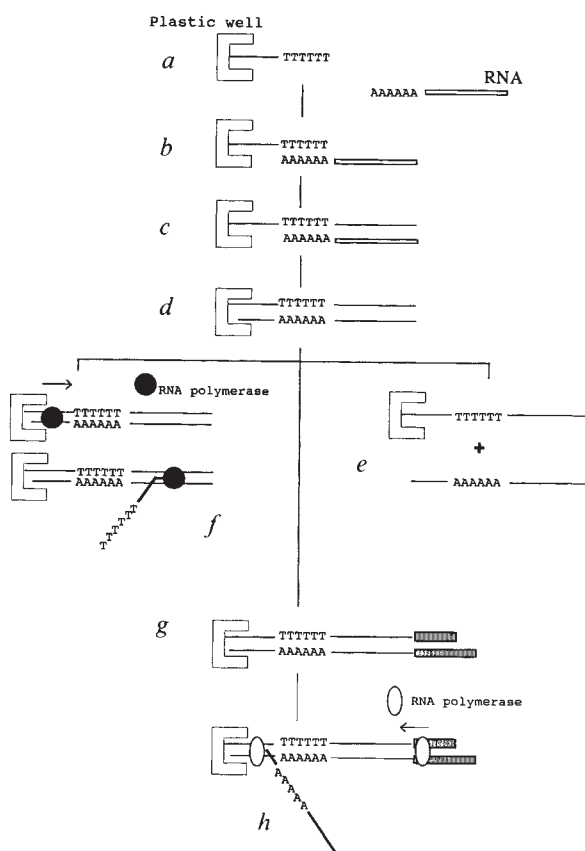


FIG. 1 Assay system based on the oligonucleotide-immobilized plastic plate. Cellular materials that are either fresh or frozen are lysed in the presence of RNase inhibitors and applied to the oligonucleotide-immobilized plastic plate (a). Poly(A)⁺ mRNA is captured on the plate (b), and then quickly processed to antisense single-stranded cDNA (c), followed by double-stranded cDNA (d). From the ds-cDNA on the plate, both strands of sense (g-h) and antisense (f) RNA can be synthesized. Furthermore, sense cDNA can be separated from antisense cDNA on the plate (e).

of a 96-well microtitre plate allow many samples to be processed on the same plate, or at the same time. Moreover, genetic information of mRNA from small quantities of invaluable clinical specimens, such as blood cells from low-birth-weight infants and biopsy materials, can be stored on the plate for long periods of time and used for the detection of specific genes at any time (Y. Tsukada, H. Mitsuhashi, K. Hiromura, T. Naruse & M. M., manuscript in preparation). It is hoped that this assay method will expand the realm of investigation for researchers in a wide range of scientific and clinical fields. □

Masato Mitsuhashi, Cylia Keller and Tatsuo Akitaya are in the Division of Medical Sciences, Hitachi Chemical Research Center, Inc., 1003 Health Sciences Road West, Irvine, California 92715, USA. For more information, fill in reader service number 100.

1. Sambrook, J., Fritsch, E. F. & Maniatis, T. in *Molecular Cloning, A Laboratory Manual: Extraction, Purification and Analysis of Messenger RNA from Eukaryotic Cells* 2nd Edn 7.1-7.87 (Cold Spring Harbor, New York, 1989).
2. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
3. Melton, D. A. et al. *Nucleic Acids Res.* **7**, 1175-1193 (1984).
4. Kawasaki, E. S. & Wang, A. M. in *PCR Technology: Detection of Gene Expression* (ed. Erlich, E. A.) 89-97 (Stockton, New York, 1989).
5. White, B. A. & Bancroft, F. C. *J. biol. Chem.* **257**, 8569-8572 (1982).
6. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning: A Practical Approach — Constructing and Screening cDNA Libraries in λ gt10 and λ gt11* (ed. Glover, D. M.) Vol. 1., 49 (IRL, Oxford, 1985).

ADVERTISEMENTS

\$3.50/Base

**Custom DNA
Synthesis**

No Set-up Fee,
Shipped in 24 hours.

**1-800-788-4362
1-800-786-4362 Fax**

**KEYSTONE
Laboratories, Inc.**

Reader Service No.33

Zeineh 1D/2D/DNA/TLC
HAND-HELD SCANNING DENSITOMETER

To quantify:

Gels, blots, TLC, X-Ray film, Photo.

1-D pkg price \$995

8 bits, 60 μ m resolution, 0-3 OD, sample 11x40 cm, IBM or clones
Options: UV, 2-D, DNA, ELISA, colony counting..... \$995 each.
Universal software: for any scanner. Main 1D, 2D, DNA, \$995 each.
Representatives needed. 30% discount of USA price list.

Do-it-all, Departmental, Laser & Video
Micro Densitometry Station
\$ 16,990



BIOMED INSTRUMENTS INC. 1020 S Raymond Av., #B, Fullerton
CA 92631. Pnn: (800) 927-2290, (714) 870-0290, Fx: 17148706385

Reader Service No.11

DNA product pipeline

What's new, what's hot? DNA sequencing tools for the Macintosh, a vector system for the selective cloning of exons and a handy hand-held scanner.

GENE CODES announces the first release of Sequencer, a **package of software tools for DNA sequencing** using the Apple Macintosh (*Reader Service No. 101*). The fast alignment algorithm in Sequencer detects overlapping ends in sequencing fragments and assembles them into a consensus sequence, commonly called a 'contig'. The program is designed to be used in assembling data generated by direct sequencing protocols as well as shotgun, random-clone sequencing. In addition, Sequencer supports transposon-facilitated sequencing with special artifact-screening functions.



Sequencer DNA sequencing program.

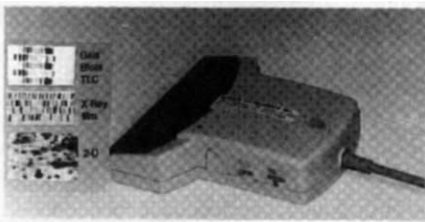
Sequencer provides a collection of quality assurance and data auditing capabilities: notes can be attached to each fragment and a copy of the original data is stored along with each 'edited sequence'. Changes to the bases in each assembly are highlighted with colour and bold characters, allowing the user to see at a glance how much the fragments were 'cleaned up' in order to fit the contig. Sequences and keystrokes can be read back audibly through the Macintosh's built-in speaker. As contamination of the GenBank and EMBL databases with vector sequence has been shown to be a serious problem (see *Nature* **355**, 211; 1992), Sequencer includes sensitive tools for detecting and eliminating vector sequences from fragment data. The program disks include a copy of VecBase, a database of commercially available cloning vectors, which was developed at the Massachusetts Institute of Technology. The artifact-screening portion of the program also addresses the new technique of sequencing with transposon insertions, where small 'tags' (four to five base duplications beyond the primer region at the site of integration) must be detected and eliminated. Other features of Sequencer include: restriction site analysis, protein translations in all frames, base counts and multiple views of assembled data, both graphically and as text. A demon-

stration disk of Sequencer is available from Gene Codes for \$10 (US) (\$15 (US) overseas), refundable with purchase.

MoBiTec offers a new **vector system for the selective cloning of exon sequences** from large genomic DNA fragments (*Reader Service No. 102*). Intron sequences are removed before cloning, considerably reducing the effort required to carry out further protocols such as sequencing the cloned DNA. Furthermore, the exontrap vector system enables the identification of a eukaryotic gene. The system can be used to derive an exon library from a genomic library of a eukaryotic chromosome or a chromosomal region, which can then be screened for specific eukaryotic genes. The exontrap function is based on a cloning vector that includes a 5' and a 3' exon separated by an intron sequence. The intron sequence contains a polylinker with many unique restriction sites. The (eukaryotic genome) DNA of interest is cloned into this polylinker. If the cloned DNA fragment contains an exon in the correct orientation, the vector expresses an RNA in which the intron sequences originating from the vector as well as those being introduced are removed; only the exons are kept. This mature RNA with only the exon sequences can then be reverse transcribed into DNA using a primer specific for the 3' exon. The DNA is amplified *in vitro* and cloned. This function is realized by virtue of a shuttle vector containing prokaryotic as well as eukaryotic genetic elements for replication (that is, the vector grows in bacteria and in animal cells in culture). The exontrap vector system includes the exontrap DNA vector, primers for cDNA synthesis and *in vitro* amplification, a handbook and protocols.

Hardware update

Measuring just 14 × 14 × 3.8 cm and weighing 0.23 kg, the new **hand-held scanning densitometer** from Biomed Instruments reads and quantitates one- and two-dimensional electrophoresis gels, blots, X-ray film (autoradiographs) and thin layer chromatography plates (*Reader Service No.*



Biomed's lightweight hand-held scanner.

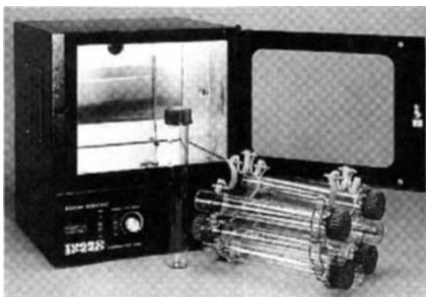
103). The \$595 (US) hand-held scanning device is intended to provide an economic alternative to other more costly high-technology instruments and provides a simple-to-use set up for scanning, analysis and hard-copy output. Special features include a resolution of 60 μm ; a reflectance scanning mode (UV, fluorescence and transmittance are optional extras); a digitizer card with an 8-bit resolution — 256 grey levels (12-bit resolution is optional); and a 3-s scanning speed for gels of 12 $\times$ 20 cm or less with 8–20 one-dimensional lanes. Optional software modules, each costing \$950 (US), include programs for scanning two-dimensional gels, DNA sequence reading/editing, colony counting, ELISA reading, image analysis, clinical electrophoresis and thin layer chromatography, antibiotic sensitivity, and microbial identification.

The ACT Model 90 **tabletop peptide synthesizer** from Advanced ChemTech is an automated dual-reactor unit that is capable of making two different peptides concurrently (*Reader Service No. 104*). It permits the use of Boc and Fmoc protocols, as well as more recent coupling agents such as HBTU and PyBOP (which do not require diimide activation) at scales of 0.1 to 10 mmols (or approximately 0.25 to 20 grams of starting resin). (ACT does not recommend use of the BOP reagent because of the carcinogenic activity of the HMPA by-product.) The unit is controlled by an optional IBM microcomputer running mouse-driven software, which includes an online 'notebook' for user comments and a printable synthesis log. Compact and self-contained, the Model 90 can accommodate seven reagent bottles and utilizes nitrogen bubbling and/or wrist action shaking for mixing. The \$16,000 (US) (\$18,000 (US) with computer) Model 90 peptide synthesizer is capable of producing four 12- to 15-mers per week.

Autogene II, the new **temperature-cycling water bath** from Grant Instruments, is designed to provide rapid temperature cycling and precise control of reactions at all times (*Reader Service No. 105*). The system's microprocessor-controlled elec-

tronics constantly monitor and adjust heat input and extraction to give an even temperature profile. This method of energy regulation ensures that each set temperature is reached in the optimum time. Up to 50 programmes can be stored in the memory, each comprising up to 99 repeats of a cycle, as well as pre- and post-treatment parameters. Programmes can also be linked, providing more flexibility by allowing parameters to be varied throughout a procedure. The system can be cooled either by the circulation of tap water or by use of Grant's optional Autocool refrigerated circulator. In addition to providing cooling, the Autocool accessory allows samples to be held at 4 °C at the end of a programme.

Hybridization ovens are rapidly becoming the method of choice among molecular biologists for Southern and northern hybridizations. However, one of the problems with hybridization ovens can be the time it takes to reach the target temperature. Bio/Can Scientific says it has overcome this problem with the Turbo-Speed **hybridization oven** by using a high capacity air turbine motor to heat and cool the interior of the oven (*Reader Service No. 106*). Whereas



Bio/Can's Turbo-Speed hybridization oven.

other ovens may take 20–40 mins to go from ambient temperature to 65 °C and equilibrate, Bio/Can says the Turbo-Speed takes 5 mins. The oven is designed to cool down quickly too — going from 65 °C to 45 °C in 5 mins. For added flexibility, the Turbo-Speed oven also features a variable speed rotator that allows the user to optimize the speed of rotation for a particular probe. Other features include a large viewing window for monitoring the run and a wide selection of hybridization bottles.

■ Kits and assorted reagents

Bios offers a **subchromosome localization kit** which enables researchers to map a polymorphic human DNA locus to a specific region of a chromosome (*Reader Service No. 107*). The kit consists of sets of DNAs derived from three-generation families represented in the Centre d'Etude du Polymorphisme Humain (CEPH) family repository. Localization is carried out by correlating the presence of a specific chromosome subfragment with the presence of a given allele. Chromosome subfragments are defined by naturally occurring physical



Eliminate hot and cold spots with Autogene II from Grant Instruments.

ADVERTISEMENT

OLIGOS

and more...

by

EXCELLENT
QUALITY:
each oligo desalted,
FPLC checked

QUICK
DELIVERY:
shipped in 2 days,
delivery by UPS

PRICES:
DM 5,-/ base
DM 95,-/ oligo
DM 195,-/ biotin

NEW :
DM 10,- per base
>1 OD FPLCpure

Oligos
Peptides
Synthetic genes
DNA sequencing

GERMANY
fax +49(30)7127734
tel +49(30)7123315

TIB
MOLBIOL

Reader Service No. 3

ADVERTISEMENTS

SINGER MSM SYSTEM
for YEAST GENETICS:-

- Ascus Dissection
- Pedigree Analysis
- Cell & Zygote Isolation

SINGER
INSTRUMENTS
"A responsibility to Science"

Call or fax now for further information
& demonstration video

Tel: +44 984 40226 Fax: +44 984 41166

Singer Instrument Co Ltd, Roadwater, Watchet, Somerset, England, TA23 0QL

Reader Service No 9

Severn Biotech Ltd.
CUSTOM SERVICES

- **OLIGONUCLEOTIDE SYNTHESIS**
High quality PROBES, PRIMERS, ANTISENSE OLIGO'S!
Fast Efficient Service offering loads of modifications,
BIOTIN, FLUORESCIN, 5' & 3' labelling,
Phosphothioates Affinity Support/Oligo's etc
- **MULTIPLE PEPTIDE SYNTHESIS**
Epitope Mapping & Screening! We now supply 5mg
quantities of high purity custom peptides up to 15
residues in length, using Fmoc chemistry. With D or L
amino acids, these are supplied lyophilised in 4ml
vials, cleaved from solid phase & ready for use.

Tel: 0562 825286 Fax: 0562 825284
Unit 23, Hookbrook Enterprise Centre
Worcester Road, Kidderminster DY10 1TY

Reader Service No 14

Professional Custom Services

DNA SYNTHESIS
PROTEIN SEQUENCING
PEPTIDE SYNTHESIS

Rapid Service • Guaranteed Quality • Ample Quantity
Competitive Prices • Free Quotations
World-wide Delivery

V.G. Services
1200 Wonderland Rd. S. Bldg. #7, Unit #2
London, Ontario, Canada N6L 1A6
CALL/FAX (519) 652-3758 ALT. FAX (519) 652-9622

Reader Service No.19

Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

£2.75 per base
(All inclusive)
0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit
King's College School of Medicine & Dentistry
Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No 56

breakpoints formed during meiotic recombination. As these meiotic breakpoints are stable, they allow Bios to generate crossover maps for each chromosome from every sibling in these families (a total of 108 individuals). These approximate 7,000 stable breakpoints identify roughly 7,000 stable fragments. Bios says that, in principle, a locus can be mapped to a fragment size that represents 0.1–1.0 per cent of the entire genome with these panels. Once a locus is assigned to a chromosome, the subchromosome localization is a two-step process. First, a polymorphism is identified for this locus in the form of a restriction length polymorphism (RFLP) (for example, by using the Bios RFLP screening kit). Second, with this information, Bios personnel construct a subchromosome localization panel specific for that polymorphism. The polymorphism is typed on this panel and Bios then assists the researcher in analysing the data using software developed at the University of Utah and at Bios. Subchromosome localization kits are available as Bios blots, Southern DNAs and PCRable DNAs. The subchromosome localization kits are provided with a set of crossover maps specific for that mapping panel and a complete protocol for carrying out these studies and analysing the data. If the subchromosome localization service is chosen, Bios will perform the laboratory procedures with a researcher's test probe or primer pair, and provide the researcher with the results and data analysis.

For the preparation of single-stranded templates from vectors containing the lacZ region, such as M13 and vectors derived from M13, Dynal recommends its new Dynabeads lacZ vector purification kit which incorporates paramagnetic Dynabeads as its key component (Reader Service No. 108). A synthetic oligonucleotide complementary to part of the lacZ region is linked to the Dynabeads through the biotin–streptavidin system (Dynabeads lacZ). Following phage lysis or denaturing of plasmid DNA, the DNA in solution is hybridized with the immobilized lacZ sequence. The Dynabeads lacZ-vector DNA complex is isolated from the solution using a strong magnet. Use of the magnet allows extensive washing of the vector DNA, providing a single-stranded template that is free of contaminants. The pure template is then eluted off the Dynabeads lacZ and is ready for applications such as cycle sequencing. Dynal says that the Dynabeads lacZ can be reused up to three times. The kit, which includes sufficient reagents for 100 purifications, consists of Dynabeads lacZ, lysis buffer, hybridization solution A and B, washing buffer, elution buffer, storage buffer and reconditioning solution.

Oncogene Science announces the availability of phosphorothioate antisense inhibition oligonucleotides, or S-oligos, to

six important human genes involved in cancer and cell proliferation (Reader Service No. 109). S-oligos are now available to p53, c-fos, c-myc, c-neu (erb-B2), c-myc and TGF-alpha. Oncogene Science says these S-oligos, which are modified oligonucleotides, are resistant to serum and cellular nucleases, are highly soluble in aqueous solutions allowing fast cellular uptake, and hybridize well to the target sequence. The S-oligos are functionally tested to ensure inhibition of translation of the target mRNA in cell cultures and are highly purified to allow simple, direct addition to cell culture media. All S-oligos are supplied with a randomized sequence of the same base composition, which can be used as a negative control. The S-oligos should be of interest to cancer researchers, neuroscientists and cell biologists.

Cruachem announces that it is offering a selection of 2'-O-Fmp protected cyanoethyl phosphoramidites, solid supports and RNA-grade ancillary reagents to facilitate the synthesis and isolation of RNA (Reader



Reagents for RNA synthesis from Cruachem.

Service No. 110). Cruachem's phosphoramidites are checked for purity by thin layer chromatography, HPLC and NMR; solvent residue is checked by gas chromatography. As a final test, a synthesis coupling efficiency test is carried out on every batch of cyanoethyl phosphoramidites before final release for sale. Cruachem also offers a smaller scale 40 nmol polystyrene-based solid support. Packed in an all polypropylene body, these DNA synthesis columns allow users to use half-strength monomers and still synthesize high quality oligonucleotides, Cruachem says. And because of the highly hydrophobic property of polystyrene, monomer solutions can also be used over a longer time once on the synthesizer. A typical crude yield for a 20-mer will be 5 O.D. units. Available in packs of 10, Cruachem's oligonucleotide purification cartridge has sufficient capacity to process the crude yield from a 0.2-µmol scale synthesis, providing a good recovery of cleaned product. Oligonucleotides are synthesized with the final trityl protecting group on. The total processing time should be about 10 mins. □